5-1-1997

Preferred lightness and chromatic image contrast reproduction

Mihai Cuciurean-Zapan

Preferred Lightness and Chromatic Image Contrast Reproduction

by

Mihai Cuciurean-Zapan

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Center for Imaging Science in the College of Science of the Rochester Institute of Technology

May 1997

Signature of the Author____________________________________

Accepted by Henry E. Rhody

Coordinator, M.S. Degree Program

Rochester Institute of Technology
Rochester, New York

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Mihai Cuciurean-Zapan has been examined and approved by the thesis committee as satisfactory for the thesis requirement for the Master of Science degree

Dr. Mark D. Fairchild, Thesis Advisor

Dr. Jonathan S. Arney

Dr. Ethan Montag

May 14, 1997
Date

THESIS RELEASE PERMISSION FORM

Rochester Institute of Technology
Center for Imaging Science

Title of Thesis: **Preferred Lightness and Chromatic Image Contrast Reproduction**

I, Mihai Cuciurean-Zapan, hereby grant permission to the Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature ______________________________

Date May 14, 1997

Preferred Lightness and Chromatic Image Contrast Reproduction

by

Mihai Cuciurean-Zapan

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Center for Imaging Science in the College of Science of the Rochester Institute of Technology

Abstract

In this study, the image preference as a function of lightness and chromatic contrast of images produced on an ink-jet printer is examined. The purpose is to develop image manipulation rules, useful in the development of printer algorithms to produce images that are preferred by viewers over images that have been printed without application of these rules. Five images are used during the psychophysical experiment, two business graphics and three pictorial, processed in three different ways in RLAB color space, once having only the lightness contrast varied, then only the chromatic contrast, and finally both lightness and chromatic contrast varied. The results showed that for the graphics images seen without a CRT original used for comparison, the mean preference was an increase in lightness contrast, while with an original available for comparison, the mean preference indicated a decrease in both lightness and chromatic contrast. For pictorial images, in the first phase of the experiment the mean preference was an increase in both lightness and chromatic contrast, and after comparison, a decrease in lightness and simultaneous decrease in lightness and chromatic contrast are the most preferred.

ACKNOWLEDGMENTS

Special thanks to my thesis advisor Dr. Mark D. Fairchild for guiding me throughout my work, for his support and encouragement.

Thanks to Dr. Jonathan Arney and Dr. Ethan Montag for their valuable help and advice.

This work was made possible by a grant from HP. I am grateful for being part of it.

Table of Contents

List of Figures

Chapter 4.

Chapter 5.

Appendix 3

List of Tables

Appendix 4

Chapter 1.

Introduction

> *"The colours are acts of light; its active and passive modifications; thus considered we may expect from them some explanation respecting light itself. Colours and light, it is true, stand in the most intimate relation to each other, but we should think of both as belonging to nature as a whole, for it is nature as a whole which manifests itself by their means in an especial manner to the sense of sight."*
>
> Goethe - "Theory of Colours", 1810.

1.1 Problem Statement and Objectives

Color is not simply a physical phenomenon dependent on the sample and illuminant. It is essentially a complex visual sensation, influenced by psychological and physiological factors that probably make one person's perception of color slightly different from another's. To understand the sensation of color, it is necessary to examine the illuminant, the characteristics of the sample, and the human factors, physiological and psychological.

There are a multitude of photometric and colorimetric tests that can be applied to a color reproduction. Among these are tests on color balance, i.e., on the absence of hue in the grays and whites, tests on chromaticity range, on the luminance range, on the contrast, on the accuracy of chromaticity reproduction of certain key colors. All of these have to be integrated in order to assess the general quality of the picture. This requires a knowledge of what makes a “good” picture.

Of course, it is very unlikely to find an answer that everyone will accept, so far as this is more as an subjective question. That means, it is necessary that a basis on which to describe the characteristics of a color picture should be found.

Compared with the monochrome case where a comprehensive assessment should include *size, definition, luminance, luminance range and luminance contrast,* the quality of a color reproduction should embrace on top of these, additional items, the most important of these being *color balance*, since the presence of any noticeable hue in the neutral grays and whites of a color picture can be objectionable[31]. The lack of neutrality is a function of the reference white against which the image is being judged. In the case of a color print, this will generally be the white border.

It is known[1] that colorimetrically correct results are not necessary for a color reproduction to be acceptable. Better images might be obtained by purposefully changing the images to produce preferred color, rather than accurate color matches. However, there is one property of the appearance of scenes that remains constant and this is the *overall color balance*. This is probably due to the physiological adaptation of the eye to the prevailing illuminant and also, partly due to the ability of observers to subconsciously discount the color of an illuminant when looking at an object, in its light.

Another attribute that has to be considered is the *chromaticity range*. The use of too vivid colors in a reproduction, most likely won't be accepted. The brilliance of some scene can not be faithfully reproduced, unless the chromaticity range is adequate. If some hues are vividly reproduced, while other are degraded, the reproduction will be unsatisfactory. A consistent mediocrity in all hues is preferred. If major hue distortions are objectionable, the small ones are usually passed unnoticed, except with very familiar objects, like sky, grass, and flesh tints.

One rule of thumb found in printers' practice and in books on the subject is the importance of contrast[1,2]. Absolute luminance levels are relatively unimportant.

Other factors that are considered, when judging a color reproduction are *gradations of hue, gradations of chroma, gradations of white content*. All these are visually significant, because saturation gradients occur frequently in nature, owing to the increasing effect of mist with distance in a landscape, to the change of surface reflection with the angle of illumination and angle of view.

If the surface texture is one of the most vital items in an aesthetically satisfying picture (for example texture of the human skin), one of the most important factors determining the ability of a color reproduction system is the color resolution of the system. If the color resolution of the system is limited, then gradations of hue and saturation, and lightness are easily lost.

As already mentioned, because the exact objective reproduction will rarely match with our preference, it is necessary to established more clearly the relationship between objective measurement and subjective appearance.

Even now, we have to agree with Evans, Hanson and Brewer[3] which almost 50 years ago mentioned that because "no simple set of colorimetric relationships between the original and the reproduction is likely to be found", some empirical approach is inevitable.

In what follows, an attempt is made, to see if based only on color preference, it is possible to derive rules for color manipulations, that applied to various images produces preferred reproductions.

1.2 Thesis Overview

Chapter 2 provides an idea about how a color reproduction is perceived, and what factors are involved in this process. A few words about hard-copy vs. soft copy, color constancy, and psychophysics terms are explained.

Chapter 3 describes the psychophysical methods necessary to interpret the results obtained during the experiment.

Chapter 4 describes the experimental environment, printer and CRT set-up, and follows the steps in obtaining the rank order of preference for an image. Also the goodness of fit of the model to the data is discussed.

Chapter 5 represents some conclusions after performing the experiment and interpreting data.

Appendix 1 contains the images used in the experiment.

Appendix 2 is a collection of plots, that suggest how the variation of lightness and chroma correlate with colorimetric quantities, measured using the

5x5x5 image target. Comparisons between original CRT CIELAB coordinates and CIELAB quantities corresponding to the prints are performed.

Appendix 3 includes the histograms of the ΔE_{94} and MCDM, and the minimum, maximum, median, and mean values.

In Appendix 4 can be found the tables including the frequencies matrices, and the rank order preference for each image.

Chapter 2.

Background

In this study, the lightness and the chromatic contrast are manipulated, to obtain images, produced on an ink-jet printer and psychophysically scaled for image preference. The aim was to provide guidance on meaningful color image manipulations that could be performed to produce printed images that are preferred over the originals.

Holding the colorimetric and color appearance conditions constant, systematically varying the exponents in the RLAB color space ($L^R a^R b^R$ exponents), an attempt is made to determine if observer preference alone, produces systematic trends. If, for a large population of observers, such trends can be found, it means that it is possible to derive rules for color image manipulations, that applied to various images produces preferred reproductions.

2.1. Hard-copy vs. soft-copy

The experiment, in the first phase, scaled images, both pictorial and graphics, for image preference (prints view in isolation). In the second phase, CRT-displayed images were used as originals for comparison.

For these reasons, it is useful to review some of the problems encountered when hard-copy versus CRT displayed images are compared, and also, what psychological and physiological factors influence our decisions when making a classification based on preference of color reproductions.

One application in the field of applied color science is that of matching hard-copy output of color printers to colored images on electronic displays. The problem is not so easy to solve, since it involves different techniques of creating color, and depends on environmental conditions. The colors on an electronic display are generated by additive mixing of the light emitted by RGB primaries, whereas the colors on the hard-copy are produced by subtractive mixing of dyes. The light reflected from a color print not only depends on the choice and amounts of dye, ink, or toner primaries that are printed on the paper, but also on the ambient light, which may vary because of changes in natural lightning, or the addition of artificial light. This is where *chromatic adaptation* becomes an important consideration, as has been recognized by the Commission Internationale d'Eclairage (CIE), which has led to the formation of a CIE Technical Committee (TC 1-27) that specifically addresses the problem of maintaining color fidelity in the process of "soft to hard-copy" color reproduction.

2.2. Color Constancy

In situations where the observer cannot or does not make a direct perceptual comparison with the original, a related phenomenon, color constancy is experienced.

Color constancy is the substantial independence of object-color perception in the presence of changes in illumination or other viewing conditions. When one looks out at the same scene on different occasions, it looks much the same as usual regardless of significant changes in the illuminant, the distance, the size, and the angle of view. Color constancy means that changes in the conditions of illumination or viewing yield no disturbing changes in the object-color perception. Though the color of the light coming from the object in the visual field is constantly changing, the perceived color does not seem to change. These remarks apply to the usual perceptions of everyday life, when there is no special effort to evaluate or question the validity of the perceptive process itself. Constancy can be greatly reduced by critical analytical scrutiny. The particular degree of color constancy actually experienced in any ordinary situation, depends greatly on several factors. Some of these are retinal and physiological, and others are judgmental or interpretive. During the experiment, the variation of the color contrast should oppose rather than favor color constancy. Color constancy also depends in part on the latitude of *memory colors* of familiar objects. Changes due to illumination may pass unremarked because they do not exceed the individual's memory color latitude or tolerance.

2.3. Factors Influencing preference

It has to be mentioned that by determining what is looked at and what is looked for, have much to do with what is actually seen or perceived. An attitude in which the observer is not so much concerned with the general nature of the object, as with the stimulation coming from that direction disfavor color constancy, and this is the attitude that the observers will be asked for, during the experiment.

All affective responses from visual stimuli must depend in some way, upon color because visual perception is impossible without some visual stimulus pattern, which in turn, is impossible without the colors that are its elements.

Aesthetic preference for some single relatively isolated color as compared with other such colors has been studied extensively by methods including paired comparisons, order of merit, and absolute judgment. Most of this work has been done with reflecting samples of different size and relatively high color purity. Not only *dominant wavelength* plays a leading role in the determination of color preferences, but also *luminance* and *purity* are significant. Due to the multiplicity of distinguishable colors, a statement of any particular color preference requires close specification in terms of all three defining dimensions. By example, affective value increases with increasing luminance, regardless of dominant wavelength, until the comfort limit is exceeded. Also, affective value increases with increasing purity up to the spectrum limit.

Another conclusion coming from the experiments testing isolated colors preference is that, affective value represents the specific position of a color in the continuum which connects the extremes of unpleasantness and pleasantness. The principle of continuity holds in the sense that neighboring colors in the color solid have similar preference.

The intensity of preferences varies because of differences in affective sensitivity to colors, as well as because of differences in the colors themselves.

Some people refer to colors by such terms as hot, juicy, sober, insipid, brutal, discordant, terms that suggest that such people pay attention to colors. Some other people pay so little attention to colors, that their typical affective responses to color can be nothing else but weak or indifferent.

Aesthetic responses are subject to *fatigue* or *adaptation*, significant losses in affective value sometimes occurring after a few seconds or minutes. Sometimes the perception itself is so changed by the sensory adaptation or *after-images*, that a different affective response is given. Affective responses such as pleasantness or unpleasantness tends to be reduced toward indifference due to pure affective adaptation.

Another factor that can influence the effective response is the *area* of a color; for example, an unpleasant color is less objectionable in small area than in large area[4,5,6].

Experiments have shown that children's preferences, develop and shift with age, moving from warm to cool colors with increasing years[7,8,9]

Preferences for color combinations as well as single colors are marked by large variations between individuals, compared to rather small variation of judgments made at different times by any particular person.

Affective color contrast enhancement is a general phenomenon that affects our preference, both in studying single colors and color combinations. The affective value of colors is raised when they follow other less pleasant or more unpleasant colors, and lowered when they follow other more pleasant or less unpleasant colors[10].

Studies of the intrinsic pleasantness of the simplest color combinations have shown that the affective value of a combination of chromatic colors is highly dependent on the affective values of the component colors. This is the principle known as the *law of affective color combination*[11,12]. The law of affective color combination appears to hold for achromatic as well as chromatic color combinations. However, the affective values of the components do not seem to be strictly additive in the color combination. The only thing that might be expected to upset would be a strong contrary effect of the combination.

When a complex color combination is the subject of the experiment, many specific factors have been suggested as effective in harmonizing the combination. The most useful percepts are[32]:

- desaturated complementaries provide the best harmony of the complementaries; a maximum of unpleasantness can result from strong or saturated complementaries in juxtaposition, due mainly to their conflicting demands on visual accommodation and the impression resulting from after-images of each color which are projected onto the neighboring color as they are fixated successively;
- regarding the hues, those that are separated by small or large hue intervals harmonize better than those separated by intervals of intermediate magnitude;
- by decreasing saturation the range of permissible hue becomes larger;
- large areas should be desaturated and conversely, the use of highly saturated colors only for small areas are desirable;

- in any composition some lightness variation is necessary to avoid monotony and provide more definition;

- excessive lightness or brightness contrast must be avoided;

- the possible danger of chromatic and achromatic contrasts and after-images should be considered.

Experimental determinations of preference have shown that in general the color in which the proportions of whiteness and blackness are in accordance with the natural lightnesses of the hue are perceived as harmonious by subjects.

If there is a figure or local region of principal interest, that region should show significant contrast with its unimportant surroundings.

Worthwhile to be mentioned is that, if optimum color relations have been worked out for a composition of one size, they are likely to require readjustments for the same composition in quite different size, because differences which are pleasing in small size tend to become too great in large size[13]

Usage of color affects preference, just as preference for color affects usage.

When the purpose of a photography/print is to provide a visual record, reproduction of colors is obviously essential even when high accuracy is not necessary. If the corresponding memory colors are matched rather than the psychophysical colors of the original scene, the greatest satisfaction is likely to result. For common natural things like human skin, sky, grass, sand, oranges, the memory colors not only seem most representative of the original colors, but are the most pleasing as well. Memory color tends to accent dominant color characteristics[14]. In portraits, significant departures from memory colors as well as original colors may be necessary to achieve maximum satisfaction.

Certain colors or color combinations are often appropriate simply because they promote the functional efficiency of the visual mechanism.

Regarding the appearance of the prints containing complex color pictures, from experience of printers and available literature, such as Yule's , *Principles of Color Reproduction*[2], the following set of rules can be deduced[33]:

1. Make the darkest achromatic color in the original the darkest achromatic color in the reproduction medium.
2. Make the lightest achromatic color in the original the most lightest achromatic color in the reproduction medium.
3. Ensure that all achromatic colors in the original remain achromatic in the reproduction.
4. Ensure that all colored regions remain in the same color category in the reproduction that they occupied in the original.

5. Make colors in the reproduction as saturated as the reproduction medium allows.

2.4. Psychophysics of reading

Most of the time, a printer is used to produce plain text. That is why, it might be useful to talk about the psychophysics of reading, mainly about how color contrast can be used to integrate form and how it interacts with luminance contrast in the task, and how color affects reading.

Probably, when talking about preference related to prints containing mostly text, a measure of the preference would be the clarity of the individual letters, which ultimately means a perfect contour, and this is reflected by the reading speed offered by the prints.

Some experiments have been designed to measure the reading speed of normal and low-vision subjects as a function of *luminance contrast, color contrast* (derived from mixtures of red and green), and combination of the two[15,16]. Normally sighted subjects showed no effect of color on reading under photopic conditions, except near the acuity limit. Only subjects with advanced photoreceptor disorders showed wavelength-specific effects in reading. The studies have shown no wavelength-specific effect in reading, and that wavelength by itself only occasionally affects reading.

Usually, letters on a background are made visible by luminance contrast (dark letters on a white background), but color contrast may play a similar role (e.g., red letters on a green background). It is possible that the color contrast may be preferable to luminance contrast for some types of low-vision reading or that color contrast in combination with luminance contrast may be beneficial.

Some experiments concluded that reading speed was governed by brightness contrast. In their article, Legge, Parish, Luebker, and Wurm [15] faced the problem of finding a metric for comparing luminance and color contrast. They constructed text by adding together red and green images, each of which had the same luminance contrast. The red and green component images were superimposed in two ways. In register they yielded yellow text, in which letters and background differed in luminance. Out of register, they yielded red-on-green text, in which the letters and the background differed in chromaticity, but had the same luminance. When the red and green component images were reduced in contrast, the in-register superposition yielded yellow text of lower luminance contrast. In the out-of-register superposition, the lower contrast component images combined to yield equiluminant text in which both red and green were mixed in the letters and background. Instead of red text on a green background (high color contrast), there was orange text on a greenish-yellow background (low color contrast).

Another idea pointed out in the above mentioned paper is the possibility that color contrast may enhance reading for text composed of large characters. This derives from the cross-over in chromatic and luminance contrast sensitivity functions at low spatial frequencies. It was found that chromatic contrast sensitivities are greater than luminance contrast sensitivities below 0.3 cycle/degree.

A problem concerning color-contrast and luminance-contrast is additivity. Once we know how reading depends on either attribute alone, we can ask how it depends on their combination. To the extent that color and luminance contrast in reading are processed in independent parallel pathways, there should be no additive interactions.

The conclusion of the experiment was that color contrast and luminance contrast act independently in their effects on reading. The lack of additivity suggested that these signals are processed in independent parallel pathways.

Some other researchers found that chromatic noise has no effect on luminance-contrast detection and that luminance noise has no effect on chromatic-contrast detection.

More recent results, Tai-Lioan Chen, and Chi-Yuang Yu [17] showed that for the OSA uniform color space, there is evidence of an additive interaction between chromatic contrast and luminance contrast.

The *OSA color space* is a color-appearance system that exemplifies uniform color spacing by means of the regular rhombohedral lattice arrangement of color samples[18] In this color space, any selected point within the specific geometrical arrangement is surrounded by twelve nearest-neighbor points, that are at an equal distance from that point.

Besides luminance contrast, our visual acuity seems to be influenced also, by chromatic contrast.

Although, whether or not chromatic contrast can affect visual performance is contradictory, more evidence is in favor of its effectiveness. Some researchers showed that chromatic contrast is unable to induce visual contrast, some indicated that chromatic contrast is capable of inducing visual contrast, but in a lower manner than luminance contrast. In the same time, in some other studies it is shown that chromatic contrast can be as effective as luminance contrast. More than that, was demonstrated that color contrast also interact with luminance contrast. Eastman Kodak Company in an report from 1944 considered that visual acuity is facilitated more by increasing luminance contrast than by increasing color contrast. Under certain conditions, if chromatic contrast is high enough, the contribution of chromatic contrast to visual performance may be as great as that of luminance contrast. For CIE color space, it was demonstrated that human performance is related to color contrast.

In OSA color space, the *luminance contrast,* ΔL is defined as the difference between the luminance of the two colors and the *color contrast* ΔC is defined by $\Delta C = \{(\Delta a)^2 + (\Delta b)^2\}^{1/2}$, where Δa is the difference between the values of green-red of the two colors, and Δb is the difference between the values of yellow-blue of the two colors.

The additive effect of luminance contrast and chromatic contrast is significant when luminance contrast is low.

According to Kirschmann's law of contrast, "color contrast is best observable when contrast of brightness is lacking or at a minimum" A good deal of experimental studies has been carried out since Kirschmann, and some of the recent experiments question the validity of this generalization. M. Alpern[19] showed that simultaneous brightness contrast is at maximum when color contrast is at a minimum.

2.5. Psychophysics of Evaluating a Scene

As already pointed out, there is no simple set of colorimetric relationships between the original and the reproduction, and when a comparison between a copy and the original has to be done, some empirical approach is inevitable. In this sense, results from psychophysics and vision experiments are most commonly used.

A scene may be examined by looking directly at the whole, or by taking it apart, examining the pieces and reassembling the whole while considering their interactions. Steps in the analytical process of viewing each unit are: *observation, analysis, interpretation and evaluation*. We can attempt these tasks directly by objective measurements or assessment, i.e., by instrumental or sensory assessment, or by connotation, in which judgments are made of the psychological responses engendered by the scene. It is inevitable that the more complex the scene, the more likely that connotative methods will be used.

The primary goal of vision is to generate descriptions of the world from the retinal images. Our descriptions are largely structured in terms of objects and surfaces, and thus one task of vision is to segment the retinal image into regions, each of which contains points imaged from one object. Image segmentation appears to be based on low-level cues, such as *continuity*, but in the same time may be based also on image *discontinuities*. In early vision, discontinuities are sensed by local oriented detectors called "edge detectors". Segmentation may occur by integrating the outputs of these local edge detectors into large contours, which delineate the boundaries of object images. Image discontinuities are caused by, among other things, the different surface properties of the objects, particularly their spectral reflectance.

Chromaticity or color may also serve for image segmentation. There are at least two ways color may be involved in image segmentation. First, regions delineated by luminance boundaries may be later grouped on the basis of a common interior color. This is established experimentally. In Ishihara tests, for example, spots are grouped together on the basis of a similar color to form recognizable shapes. Furthermore, the color of texture elements has been shown to mask other texture segregations, e.g. that based on orientation. Some models of perception use color this way, relegating it to merely filling-in regions delineated by luminance defined boundaries. The second way color could be used is at the region boundaries themselves. If each local edge is also sensed by a color-contrast themselves. If each local edge is also sensed by a color-contrast system, the colors on each side of the edge could assist in the grouping of the edges into boundaries.

The structure of the scene may be described in terms of molecules arranged in particular geometries in space. The structure related to its environment provide the stimulus. The stimulus is modified by retinal and neural characteristics into the appearance response, which can be defined by critical analysis. Appearance response is converted, via the viewer's temperamental factors, into consumer images, quality judgments, and preferences. Although the structure and the stimulus may be physically handled and measured, the later stages in the above sequence cannot. There is some understanding of the neural and temperamental factors associated with appearance and texture. However, not enough is known about them to allow safe prediction of preference to be made from a physical specification of structure.

A better understanding of images resulting from a scene can be gained from a more detailed look at the factors affecting *total appearance*.

2.6. Recommendations in Evaluating a Color Reproduction

Influencing the progress toward the present concept was the independent search initiated by the Inter-Society Color Council Project Committee 33. The

Committee comprised a list of factors believed to influence the human response to color. These factors were formulated into a logical model in which perceived appearance images could be derived. Such images are built from the lighting of the scene, physically definable object properties, the perceiver's inherited and learned responses, and the perceiver's immediate environmental factors.

Each perception involves memories, concepts, and attitudes. These in turn are affected by the way we feel.

Perception of color and appearance are unique to the individual. They change with color vision abilities, the state of visual adaptation, color contrast, after-image, color constancy, discrimination and metamerism characteristics.

Using standard methodology, color can be specified instrumentally, or quantified by matching to a color atlas. Visual judgments can be made using for example the elements of Natural Color System (NCS) or color appearance methodology. Within color appearance methodology, Hunt defines three attributes: brightness, hue, and colorfulness. In the real world of light and shades, and indoor and outdoor scenes, in which different intensities and qualities of illumination are present, it is necessary for each area under examination to also define the three relative subjective terms: lightness, chroma, and saturation.

Sometimes in the analysis of sensory responses to images, use can be found for sophisticated mathematical techniques such as multidimensional scaling, and principal component analysis. This type of analysis has been applied successfully to color perception.

The psychophysical evaluation of color appearance is quite complex. That's why some CIE Technical Committees were created in order to make baseline recommendations on what variables should be used in the design of the experiments and how each variable should be treated, recommendations regarding viewing conditions, illumination conditions, background and surround conditions, types of stimuli to be used, and issues on viewing techniques.

The most used color spaces in experiments on color-appearance are CIELAB and CIELUV But these spaces have no mechanism for predicting appearance changes due to change in luminance level, background or surround. So, requirement for new color-spaces that take into account these parameters led to the creation of some modifications of CIELAB model, like LABHNU and the Fairchild and Berns model[20], called RLAB. Prior to the RLAB color space, Fairchild[21] proposed a new model for chromatic adaptation. The aim of the new model was to be psychologically plausible, simple, colorimetrically accurate and to account for incomplete chromatic adaptation. However, the model has two limitations. The first one is that the change in perceived image contrast for different surround conditions is not addressed. The second one is that the model can be used to calculate corresponding colors, but not color appearance, no color-appearance metric being defined.

RLAB color space represents a combination of the model of incomplete chromatic adaptation, defined by Fairchild , the CIELAB color space, and extensions to account for changes in surround relative luminance.

This color space fits perfectly for the needs of the experiment in many respects. The illuminant required in our experiment is the same one as the reference illuminant of the RLAB color space, i.e., CIE Illuminant D65. The output images will be manipulated in lightness and chromatic contrast by varying the exponents in the RLAB color space. In the RLAB color appearance model, the contrast levels are defined by variable exponents that are specified according to the relative luminance of the surround. Previous psychophysical experiments have suggested that these exponents depend extremely on the viewing conditions, and that the optimum exponents, might depend on observer preference.

Of primary importance in visual experiments is the color and illuminance levels of the illumination used. Spectral power measurements of the light sources and the evaluation of the stability of viewing booth should be performed. The illumination level of the light source should be specified either as illuminance in lux or as the luminance of a reference white in cd/m^2. Also, the color of the light source either in xy or $u'v'$ chromaticity coordinates, should be specified.

According to CIE TC 1-34, the CIE 1931 Standard Colorimetric Observer (2°) should be used for all colorimetry.

Data must be collected for a fairly large number of observers to provide the statistical variability required for analysis. In our case at least 30 observers will be used.

Either a magnitude scaling method can be used, in which observers assign a scale value to their preference, or a rank-order procedure in which the observers rank their preference. Although, these techniques will provide useful data, the scales are not likely to be as precise as an interval scale derived via the method of paired comparisons, is the conclusion of the CIE TC 1-34.

The method of paired-comparison is an elaboration of the method where one stimulus serves as a standard for comparison with the other stimuli in the series. Fechner conceived the idea that a psychophysical experiment could be conducted in which an observer makes judgments on a psychological dimension having no obvious physical correlate. In his book on the experimental study of aesthetics, he suggested that the pleasantness of two objects could be studied by having observers choose the object that was more pleasant. The first experimental study in which this method was employed was an investigation of color preference. A theoretical analysis of the type of data provided by the method came in 1927, when Thurstone published his paper on the law of comparative judgment as applied to paired comparison judgments.

The method of paired-comparison is the method most frequently employed to collect data for constructing psychological scales based upon comparative judgments. In this method the observer is required to make comparative judgments for all possible pairs of stimuli.

Preference frequency data collected via paired-comparison tests can be converted into interval scales using Thurstone's Law of Comparative Judgments.

If the frequency data is derived from the proportion of times each stimulus was judged to be in each category of a set of categories which are ordered with respect to a given attribute, than the data is converted into interval scale using the "law of categorical judgment", which is discussed in Chapter 3.

Chapter 3.

Psychophysics

3-1. The Law of Categorical Judgment

The "law of categorical judgment" is a set of equations relating parameters of stimuli and category boundaries to a set of cumulative proportions derived from the proportions of times each stimulus is judged to be in each category of a set of categories which are ordered with respect to a given attribute. Like the law of comparative judgment it is based on Thurstone's general judgment model. They can be summarized as follows :

> "A psychological continuum of attribute of interest is postulated. Each time a stimulus is presented to a subject, it brings about the same sort of a discriminal process which has a value on this continuum. Owing to various factors, upon repeated presentation, this stimulus is not always associated with a particular value, but may be associated with one higher or lower on the continuum. It is postulated that the values associated with any given stimulus project a normal distribution on the continuum. Different stimuli may have different means (scale values) and different deviations (discriminal dispersions)."
>
> Torgerson, W.S.

To derive the law of categorical judgment Torgerson used the following assumptions

- The psychological continuum of the subject can be divided into a specified number of ordered categories or steps.
- Owing to various factors a given category boundary is not necessarily always located at a particular point on the continuum. Rather, it also projects a normal distribution of positions on the continuum. Again, different category boundaries may have different mean locations and different dispersions.
- The subject judges a given stimulus to be below a given category boundary whenever the value of the stimulus on the continuum is less than that of the category boundary.

The set of equations describing the law of categorical judgment are

(1) $t_g - s_j = z_{jg}(\sigma_j^2 + \sigma_g^2 - 2r_{jg}\sigma_j\sigma_g)^{1/2}$ $\quad j=1,2,\ldots,n$

$g=1,2,\ldots,m$

where the terms are define as follows:

- m+1 - number of categories
- t_g - mean location of the gth category boundary
- σ_g - dispersion of the gth category boundary

- r_{jg} - correlation between momentary positions of stimulus j and category boundary g
- z_{jg} - unit normal deviate corresponding to the proportion of times stimulus j is sorted below boundary g.

This equation represents the complete form of the law of categorical judgment. It is not solvable in its complete form, since there are always more unknowns than equations. So, simplifying hypotheses are necessary.

Based on different ways used to generate the proportions and the special conditions resulting from various simplifying assumptions the following classes of models are used:

- *Class I* : models involving replication over trials within a single subject.
- *Class II* : models involving replication over individuals, each stimulus being judged once by each subject.
- *Class III* : mixed models, involving replication over both individuals and trials.

Several sets of restrictions which lead to solvable versions of the general equations might be developed. The most commonly used are:

- *Condition A*. It is assumed that the covariance term of the equation (1) is constant over all values of j and g. The equation (1) reduces to

(2) $t_g - s_j = z_{jg}(a_j^2 + b_g^2)^{1/2}$, j=1,2,...,n

g=1,2,...m

where

$$a_j = (\sigma_j^2 - c)^{1/2}$$
$$b_g = (\sigma_g^2 - c)^{1/2}$$

For practical purposes the assumption is that r_{jg} is zero, this being the only reasonable way for the covariance to be constant when the variances vary. Condition A is solvable only in theory.

- *Condition B*. Under the assumption that σ_g is constant for all values of g and the correlation term vanishes, equation (1) reduces to

(3) $t_g - s_j = z_{jg} a_j$, j=1,2,...,n

g=1,2,...,m

where

$$a_j^2 = \sigma_j^2 + c.$$

Equation (3) represents the equation underlying the general method of successive intervals. This is the scenario that best fits to the case where the proportions are obtained from responses of a group of subjects.

- *Condition C*. The σ_j is constant for all stimuli ($\sigma_j^2=k$) and all correlation terms vanish. Defining $b_g^2 = \sigma_g^2 + k$, the equation (1) reduces to

(4) $t_g - s_j = z_{jg} b_g$, $j=1,2,\ldots,n$

$g=1,2,\ldots,m.$

- *Condition D*. It is assumed that σ_j is constant, σ_g is constant, and r_{jg} is constant for all values of j and g. The new obtained equation is

(5) $t_g - s_j = z_{jg} c$, $j=1,2,\ldots,n$

$g=1,2,\ldots,m.$

3-2.The Method of Successive Intervals

With respect to the experimental operations of obtaining judgments, the method of successive categories is a very general one. In principle it includes what has been called the *method of single stimuli* as well as all *rating methods* in which categorical judgments are made.

The critical assumption to scaling stimuli by means of judgments in successive categories is that the distribution of responses to a stimulus is normal on psychological continuum.

If we were to assume that the categories do actually represent equal psychological intervals, the frequency distributions of simuli are obviously not normal. Some distributions would be positively , some negatively skewed. Skewing depends on whether the mean is on the upper or lower side of the middle category, and more fundamentally, when is about data obtained scaling stimuli by means of judgments in successive categories, skewing is a function of inequality of scale units. It might happen that the category widths increase/decrease systematically as we go up the scale. For these reasons, it is necessary to give to the categories a proper width, and this is done by the scaling processes.

There are a number of ways to do this. Regardless of the particular experimental method used, the immediate raw data is in the form of the frequency with which each stimulus is rated into each category.

Chapter 4.

4-1. Discussion of the Experiment

- *Image Sources*

In this study both pictorial and graphics images were used. Three pictorial images named, "Building", "Fruits", and "Musicians", and two graphics images, that are referred to as, "Landscape Graphic", and "Seed Graphic", have been used (see Appendix 1). All these images have a resolution of 150 pixels/inch, and their size is respectively, 6"x 9" for "Building", 9" x 6" for "Fruits" and "Musicians", 9.41 " x 8 " for "Landscape", 6.87 " x 10" for "Seed"

- *Image Output*

All images used have a full-page U.S. letter size, and are produced on an HP DeskJet 870Cxi ink-jet printer, on high quality HP Premium glossy paper. The prints were appropriately mounted for completion of the visual experiments. The output images were manipulated in lightness and chromatic contrast by varying the exponents in the RLAB color space. The manipulations were made in three blocks. The first block includes 7 different levels of lightness contrast (L^R exponent) at constant chromatic contrast. This means, that for the lightness, seven values for the exponent were used; -1.5, -1.3, -1.1, 0, +1.1, +1.3, +1.5, where "-"

indicates division of the exponent by that number, "+", represents a multiplication of the exponent by the corresponding number, and 0 means no change. The second includes 7 different levels of chromatic contrast ($a^R b^R$ exponents) at a constant lightness contrast (the same values as before). The third includes the same 7 different contrast levels with both lightness and chromatic contrast covarying ($L^R a^R b^R$ exponents).

A complete description of RLAB calculations can be found in Fairchild and Berns[29] Here, only the formulas for L^R, a^R, and b^R are given in order to understand how the lightness and chromatic contrast were varied.

(1) $L^R = 116(Y/100.00)^{k\sigma} - 16$

(2) $a^R = 500[(X/95.05)^{k\sigma} - (Y/100.00)^{k\sigma}]$

(3) $b^R = 200[(Y/100.00)^{k\sigma} - (Z/108.88)^{k\sigma}]$.

The denominators in Eqs. (1)-(3) are the tristimulus values of CIE Illuminant D65. When k is in the set A={1.1, 1.3, 1.5}, the corresponding RLAB coordinates are increased, and for k in the set B={1/1.1, 1/1.3, 1/1.5}, the corresponding values are decreased. When only the lightness is varied, the corresponding images will be denoted by L*k*m, if k belongs to the set A, or L*k*d, if k belongs to the set B. To indicate the variation of the chromatic contrast, AB*k*m and AB*k*d are used, and for both lightness and chromatic contrast, the notation is LAB*k*m and LAB*k*d.

- *Experimental Environment*

A viewing environment was constructed with the appropriate illumination color and luminance level. The room was divided in two, by a black curtain to prevent light from falling on the monitor when the observer had to compare the image with the original displayed on the monitor. Also, to prevent the light from reflecting back into the observer's eyes, the wall behind the monitor was painted black. The other three walls were neutral gray.

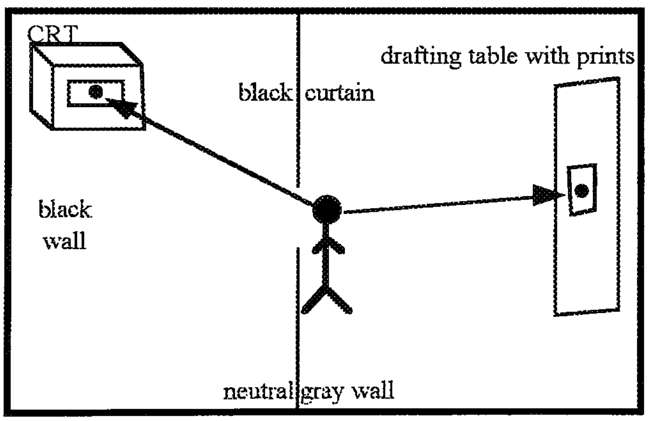


Figure 4-1-1. Top view of setup used for the experiment

The illuminant used for viewing the prints was an approximation of CIE Illuminant D65, having CIE 1931 chromaticity of x=0.306 and y=0.321, and a CCT of 6900 K.

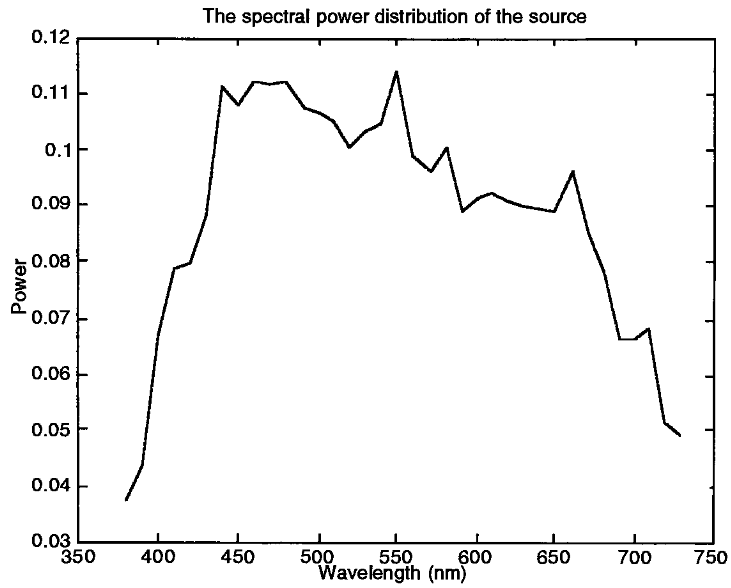


Figure 4-1-2. The illuminant source.

The setup was designed to allow viewing of printed and CRT-displayed images of equal size at equal viewing distances, approximately 50 cm. The observer could view only one copy at a time, but could switch between them at will, by rotating in his chair.

The luminance level and white-point of the CRT-display and print illumination were made to be almost equal, to avoid confounds of appearance modeling that are caused by white-point and/or luminance changes. This allowed direct comparison of the CRT and printed images with no concern about undefined chromatic or light adaptation levels. The most appropriate white-point is that of CIE illuminate D65 since it appears achromatic in both hard-copy and soft-copy displays over a range of luminance levels. The CRT display was a SGI Sony GDM-2000TC, having a display gamma of 2.2, and the white point x=0.317, y=0.333. The luminance was set on CRT display to 80 cd/m^2.

The white-point for the CRT used in calculating CIELAB coordinates had the following XYZ tristimulus values : Xn=95.57, Yn=100, and Zn=106.29.

The monitor was calibrated using the "gog-model" developed by Berns et. al.[30] The normalized matrix obtained during calibration, necessary to obtain the XYZ device-independent coordinates is

$$M = \begin{bmatrix} 47.20 & 30.95 & 17.41 \\ 25.92 & 65.97 & 8.11 \\ 2.65 & 11.74 & 91.89 \end{bmatrix}.$$

The parameters of the "gog-model" for the red, green, and blue channel are respectively:

- red - gain = 1.02238, offset = -0.01850, gamma = 1.95431;
- green - gain = 1.00592, offset = -0.00353, gamma = 2.03438;
- blue - gain = 0.96559, offset = -0.03573, gamma = 2.18274.

- *Characterize Printer*

The prints had been obtained using a HP DeskJet 870Cxi printer, having a color resolution of 600x300 dpi. In order to represent typical usage, the default ColorSmart setup was used. The accuracy of the characterization was evaluated.

A 5x5x5 Target was used to colorimetrically evaluate the printer's behavior versus the CRT image. This also allows analysis of gamut mapping and the effects of the image manipulations. A set of 5 prints of the target was averaged to obtain a mean spectral reflectance, and this average was used in our colorimetric calculations, as compared to the CRT values for the target.

The results of the comparison can be found in Appendix 2..

- *Preference Scaling Experiment 1*

23 variations of each image were scaled (18 different processed images, and 5 originals). A category rating experiment was completed in order to derive interval scales of image preference as a function of the various lightness and chromatic contrast defining exponents. A population of 34 observers, 6 experts and 28 non-experts (tested for normal color vision) was used to provide the required statistical significance. Because no significant difference was noted between these two groups, the results were pooled together. A similar analysis was completed for changes in image type and content. The first experiment was carried out by presenting printed images to the observers, one at a time, with no reference to an original image.

Before starting the experiment, the observers were instructed to classify images in five categories, "Poor", "Fair", "Good", "Excellent", and "Ideal", based only on their preference of color reproduction.

- *Preference Scaling Experiment 2*

The second experiment was completed using the same techniques as the first experiment. However, in this case, observers were making their preference judgments of the prints relative to an original image presented on a CRT display.

- *Interpretation of Results*

Once the visual data were collected and analyzed, the results were evaluated to determine whether general image color manipulation rules can be derived for the production of preferred images. The idea was that if such rules can be derived, a follow-up experiment is to be conducted to verify that the suggested rules, do indeed produce images that are preferred over those produced with no color manipulation. The rules might also provide useful guidance in the development of gamut-mapping algorithms.

An example of processing the visual data is presented in the next paragraph.

4-2. Example

As an example we'll use the data corresponding to the image "Building", when no original was available.

For the other images, the results are in Appendix 4.

Table 4-2-1. The Raw Frequency Matrix **F**

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	5	6	17	6	0
L1.1d	3	9	12	8	2
L1.3m	2	7	8	7	10
L1.3d	0	4	11	14	5
L1.5m	1	3	7	12	11
L1.5d	0	2	3	12	17
AB1.1m	0	2	12	18	2
AB1.1d	0	2	11	17	4
AB1.3m	0	4	18	10	2
AB1.3d	0	1	4	16	13
AB1.5m	2	6	10	15	1
AB1.5d	0	0	3	8	23
LAB1.1m	5	8	11	10	0
LAB1.1d	2	2	19	9	2
LAB1.3m	3	4	11	9	7
LAB1.3d	1	3	4	16	10
LAB1.5m	0	2	4	7	21
LAB1.5d	0	2	3	5	24
ORIGINAL	2	3	16	10	3

The next matrix that is constructed is the so-called *the cumulative frequency matrix*, whose elements (j,g) are equal to the number of times stimulus j was sorted below the gth category limit. The gth category boundary is being defined as the upper boundary of the gth category.

Table 4-2-2. The Cumulative Frequency Matrix Φ

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	5	11	28	34	34
L1.1d	3	12	24	32	34
L1.3m	2	9	17	24	34
L1.3d	0	4	15	29	34
L1.5m	1	4	11	23	34
L1.5d	0	4	7	17	34
AB1.1m	0	2	14	32	34
AB1.1d	0	2	13	30	34
AB1.3m	0	4	22	32	34
AB1.3d	0	1	5	21	34
AB1.5m	2	8	18	33	34
AB1.5d	0	0	4	11	34
LAB1.1m	5	13	24	34	34
LAB1.1d	2	4	23	32	34
LAB1.3m	3	7	18	27	34
LAB1.3d	1	4	8	24	34
LAB1.5m	0	4	10	16	34
LAB1.5d	0	3	6	10	34
ORIGINAL	2	5	21	31	34

The following matrix represents the proportion of times stimulus *j* was judged to be below the *g*th category boundary.

Table 4-2-3. The Cumulative Proportion Matrix **P**

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	0.15	0.32	0.82	1	1
L1.1d	0.09	0.35	0.71	0.94	1
L1.3m	0.06	0.26	0.5	0.71	1
L1.3d	0	0.12	0.44	0.85	1
L1.5m	0.03	0.12	0.32	0.68	1
L1.5d	0	0.12	0.21	0.5	1
AB1.1m	0	0.06	0.41	0.94	1
AB1.1d	0	0.06	0.38	0.88	1
AB1.3m	0	0.12	0.65	0.94	1
AB1.3d	0	0.03	0.15	0.62	1
AB1.5m	0.06	0.23	0.53	0.97	1
AB1.5d	0	0	0.12	0.32	1

LAB1.1m	0.15	0.38	0.71	1	1
LAB1.1d	0.06	0.12	0.68	0.94	1
LAB1.3m	0.09	0.21	0.53	0.79	1
LAB1.3d	0.03	0.12	0.24	0.71	1
LAB1.5m	0	0.12	0.29	0.47	1
LAB1.5d	0	0.09	0.18	0.29	1
ORIGINAL	0.06	0.15	0.62	0.91	1

The last matrix that will be used is called *the basic transformation matrix*, and has the elements representing the unit normal deviates corresponding to the elements of matrix **P**.

Table 4-2-4. The Basic Transformation Matrix **X**

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR
L1.1m	-1.05	-0.46	0.93	
L1.1d	-1.35	-0.38	0.54	1.56
L1.3m	-1.56	-0.63	0	0.54
L1.3d		-1.19	-0.1	1.05
L1.5m	-1.89	-1.19	-0.5	0.46
L1.5d		-1.19	-0.8	0
AB1.1m		-1.56	-0.2	1.56
AB1.1d		-1.56	-0.3	1.19
AB1.3m		-1.19	0.38	1.56
AB1.3d		-1.89	-1	0.3
AB1.5m	-1.56	-0.72	0.07	1.89
AB1.5d			-1.2	-0.5
LAB1.1m	-1.05	-0.3	0.54	
LAB1.1d	-1.56	-1.19	0.46	1.56
LAB1.3m	-1.35	-0.82	0.07	0.82
LAB1.3d	-1.89	-1.19	-0.7	0.54
LAB1.5m		-1.1	-0.5	-0.1
LAB1.5d		-1.35	-0.9	-0.5
ORIGINAL	-1.56	-1.05	0.3	1.35

Note: Any cells of the matrix **P** that contain proportions of zero or unity cannot be transform into z values, and therefore the cells of matrix **X** corresponding to such cells must be left vacant.

Matrix **X** contains the empirical estimates z'_{jg} of the equations of the law of categorical judgment. Because with experimentally obtained data, the z values will be in error, a number of procedures have been devised to obtain estimates from fallible data of the scale values and discriminal dispersions of the stimuli, and of the locations of the category boundaries. But all solutions have in common the four matrices described above.

Some solutions require that the matrix **X** is complete, i.e., has no vacant cells. If vacant cells are present in but a few rows of the matrix, the method can be used if these rows are omitted from calculations. However, few if any of the stimuli will be present in all categories, a different method has to be used.

In our case, having 19 stimuli (19 rows in each matrix), we can afford to omit some rows. Most of the time the omitted rows correspond to the stimuli that have their mean shifted towards the "Poor" category. For these stimuli, anyway the theory does not apply, because they are not normally distributed, and by eliminating them we do not loose any information, an elimination automatically correlates with the idea that that image is the least preferred.

The next step in our approach is to scale the category limits. One general principle of successive-categories scaling is to determine values for the limits of the categories. This limits are essentially threshold values. Another general principle is to determine a single scale value for each category.

It has been assumed that the frequency distribution of judgments of each stimulus is normal on an interval scale. The cumulative proportions given in Table 4-2-3, are taken to represent the areas under the unit normal distribution curve below the upper limits of the respective category intervals. The linear distances of those limits from the means of the stimuli, are found by looking up the corresponding deviates in the tables of the normal distribution (see Table 4-2-4). Each element in matrix **X** may be regarded as the distance of an upper category limit from the mean for that stimulus. The means for different stimuli will vary. There are also differences in dispersions, some of them due to sampling errors. Because of these, the deviates in any one column are not equal. We have as many scales as there are stimuli, each with its own unit and origin. From all this information we have to extract a single set of values for the upper limits. There is a possibility of evaluating every limit except the upper one for category "Poor" and the lower one for category "Ideal". These are unscalable because the corresponding proportions are 1 and 0, respectively, and the deviates are infinite.

In order to reduce the number of the estimates of a threshold to a single value, some kind of averaging is employed. If it is assumed that the dispersions of the stimuli are equal except for sampling errors, it is obviously justifiable to average results from different distributions. If the **X** matrix is complete, the means are found simply by dividing the sum on each column, by the number of stimuli. If the matrix is incomplete, as many estimates as is possible of category widths are determined by subtracting the deviates by pairs down neighboring pairs or columns.

Table 4-2-5.

STIMULUS	EXCELLENT	GOOD	FAIR
L1.1m	0.6	1.39	
L1.1d	0.97	0.92	1.02
L1.3m	0.93	0.63	0.54
L1.3d		1.04	1.2
L1.5m	0.7	0.73	0.92
L1.5d		0.37	0.82
AB1.1m		1.34	1.79
AB1.1d		1.27	1.49
AB1.3m		1.56	1.19
AB1.3d		0.84	1.35
AB1.5m	0.84	0.8	1.82
LAB1.1m	0.75	0.84	
LAB1.1d	0.38	1.64	1.11
LAB1.3m	0.53	0.89	0.75
LAB1.3d	0.7	0.47	1.26
LAB1.5m		0.65	0.47
LAB1.5d		0.42	0.39
ORIGINAL	0.52	1.35	1.05
SUM	6.92	17.1	17.1
Md=w	0.69	0.95	1.07
Lc=cw	0.69	1.64	2.72
Mc	0.35	1.17	2.18
Ac	-0.82	0	1.01

Here

Sum is the sum on columns,

Md=w gives us the average estimates of category widths,

Lc=cw provide us the scale values of the upper limits of the intervals,

Mc is midpoint of category limits, and

Ac is the absolute category value with zero at midpoint indifference category.

The next step is to determine scale values and variability for stimuli.

Now any of the threshold scales we have, based upon the values *Lc, Mc*, or *Ac*, can be used to assign numerical values to the stimuli. The first scale which utilizes the thresholds *Lc*, is well adapted to the computation of the medians of the stimuli. The other two, based on *Mc* and *Ac* are better adapted to the computations of means for stimuli.

There are severe limitations to the possibility of computation of medians, due to the markedly truncation of some distributions. Truncation does not preclude the computation of the median unless more than 50 per cent of the frequencies fall in an end category. Another problem is that it is impossible to assign values to judgments falling in the end categories. The interpolated medians will be used, their values being obtained using the formula[28]

$$\tilde{x} = \tilde{L} + b\left(\frac{n/2 - (\Sigma f)_{\tilde{L}}}{f_{Median}}\right) \quad (6)$$

, where

$\tilde{L}$ = lower limit of the median class,

b = class width,

n = number of values,

$(\Sigma f)_{\tilde{L}}$ = sum of the frequencies classes below the median class,

f_{Median} = number of values in median class.

Table 4-2-6. Interpolated Median values

STIMULUS	Median
L1.1m	1.028
L1.1d	1.089
L1.3m	1.644
L1.3d	1.798
L1.5m	2.18
L1.5d	2.716
AB1.1m	1.823
AB1.1d	1.897
AB1.3m	1.38
AB1.3d	2.448
AB1.5m	1.549
LAB1.1m	1.039
LAB1.1d	1.344
LAB1.3m	1.667
LAB1.3d	2.247
ORIGINAL	1.406

It is possible to estimate standard deviations for stimuli, on the psychological scale. If we plot the deviate values for each stimulus, as given in Table 4-2-4, as a function of the threshold values *Lc*, we obtain straight line regressions. The fact that each regression line is linear on the threshold scale is evidence of normality of the distribution for the stimuli. The slope of each regression line is the reciprocal of the standard deviation for the stimulus.

Table 4-2-7.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	SLOPE	StdDev
L1.1m	-1.05	-0.46	0.93		1.46	0.69
L1.1d	-1.35	-0.38	0.54	1.56	0.96	1.04
L1.3m	-1.56	-0.63	0	0.54	0.58	1.73
L1.3d		-1.19	-0.1	1.05	1.11	0.9
L1.5m	-1.89	-1.19	-0.5	0.46	0.81	1.23
L1.5d		-1.19	-0.8	0	0.59	1.69
AB1.1m		-1.56	-0.2	1.56	1.55	0.65
AB1.1d		-1.56	-0.3	1.19	1.36	0.74
AB1.3m		-1.19	0.38	1.56	1.35	0.74
AB1.3d		-1.89	- 1	0.3	1.09	0.92
AB1.5m	-1.56	-0.72	0.07	1.89	1.3	0.77
LAB1.1m	-1.05	-0.3	0.54		0.88	1.13
LAB1.1d	-1.56	-1.19	0.46	1.56	0.81	1.24
LAB1.3m	-1.35	-0.82	0.07	0.82	0.86	1.16
LAB1.3d	-1.89	-1.19	-0.7	0.54	1.18	0.85
ORIGINAL	-1.56	-1.05	0.3	1.35	1	1
Lc=cw		0.7	1.64	2.72		

The following table contains the final estimated values for the median of the stimuli and their standard deviation, ordered by increasing median value.

Table 4-2-8.

STIMULUS	Median	StdDev
L1.1m	1.03	0.69
LAB1.1m	1.04	1.13
L1.1d	1.09	1.04
LAB1.1d	1.34	1.24
AB1.3m	1.38	0.74
ORIGINAL	1.41	1
AB1.5m	1.55	0.77
L1.3m	1.64	1.73
LAB1.3m	1.67	1.16
L1.3d	1.8	0.9
AB1.1m	1.82	0.65
AB1.1d	1.9	0.74
L1.5m	2.18	1.23
LAB1.3d	2.25	0.85
AB1.3d	2.45	0.92
L1.5d	2.72	1.69
LAB1.5m		
LAB1.5d		
AB1.5d		

The last three rows have been excluded during the classification, because of the frequencies that are higher toward the "Poor" category. For this reason either in matrix **X** the first columns corresponding to these stimuli are empty, or when the interpolated median is calculated, the value of the upper boundary of the corresponding median class is not finite, so the interpolated median can not be calculated for those stimuli. Therefore, these stimuli have to be classified in some other way. A rigorous way would be to perform a two-tailed test, in order to see which mean value is lower. But, due to the fact that these happens for the stimuli which are least preferred, and reduced number of categories, nothing is lost if a direct histogram comparison is perform.

Using a linear transformation, the individual scales have been normalized, such that for all images the median values are normally distributed, with zero mean, and standard deviation equal with 1.

Table 4-2-9. Normalized median values

STIMULUS	Normalized median
L1.1m	-1.34
LAB1.1m	-1.32
L1.1d	-1.22
LAB1.1d	-0.72
AB1.3m	-0.64
ORIGINAL	-0.59
AB1.5m	-0.3
L1.3m	-0.12
LAB1.3m	-0.07
L1.3d	0.19
AB1.1m	0.24
AB1.1d	0.38
L1.5m	0.95
LAB1.3d	1.08
AB1.3d	1.48
L1.5d	2.01
LAB1.5m	
LAB1.5d	
AB1.5d	

Using the normalized data, the 95% confidence intervals are given by $\mp 1.96 \frac{1}{\sqrt{34}} = \mp 0.3361$.

In order to test the observers' consistency, in each set of images, besides the processed ones, 5 originals were included. The confidence intervals corresponding to these stimuli are given in the next table:

CONFIDENCE INTERVALS FOR THE ORIGINALS		
BUILDING	no original	0.02956
BUILDING	with original	0.02768
FRUITS	no original	0.02306
FRUITS	with original	0.03269
LANDSCAPE	no original	0.02263
LANDSCAPE	with original	0.03148
MUSICIANS	no original	0.02241
MUSICIANS	with original	0.02241
SEED	no original	0.02623
SEED	with original	0.02509

Table 4-2-10. Confidence intervals for the originals

To obtain information about the variation of the standard deviations, a statistical test of homogeneity was perform and it can be concluded that the variations are after all a matter of sampling fluctuations.

Chi-square values for the test of homogeneity of variances estimated for the images		
IMAGE	**chi-square**	**Degrees of freedom**
BUILDING - no original	93.81036766	15
BUILDING - with original	51.48441951	15
FRUITS - no original	44.15349537	14
FRUITS - with original	96.70446214	17
LANDSCAPE - no original	115.0480409	19
LANDSCAPE - with original	43.41124093	17
MUSICIANS - no original	29.61077661	13
MUSICIANS - with original	56.24217433	16
SEED - no original	77.87749527	16
SEED - with original	102.5145309	16

Table 4-2-11. Chi-square values

4-3. Goodness of Fit of the Model to the Data

If the matrix **X** is known, which means the proportion of times each stimulus is judged below each category boundary is know, the relevant parameters of the stimuli and category boundaries can be estimated. In evaluating goodness of fit, an easy check on the model is to plot the rows of matrix **X** against each other. Systematic departures from linearity would indicate that the assumptions underlying the procedure used have not met. This is true for *Condition B* from the law of categorical judgments.

As can be seen in our case, we have a good fit of the model to the data. Here are some plots:

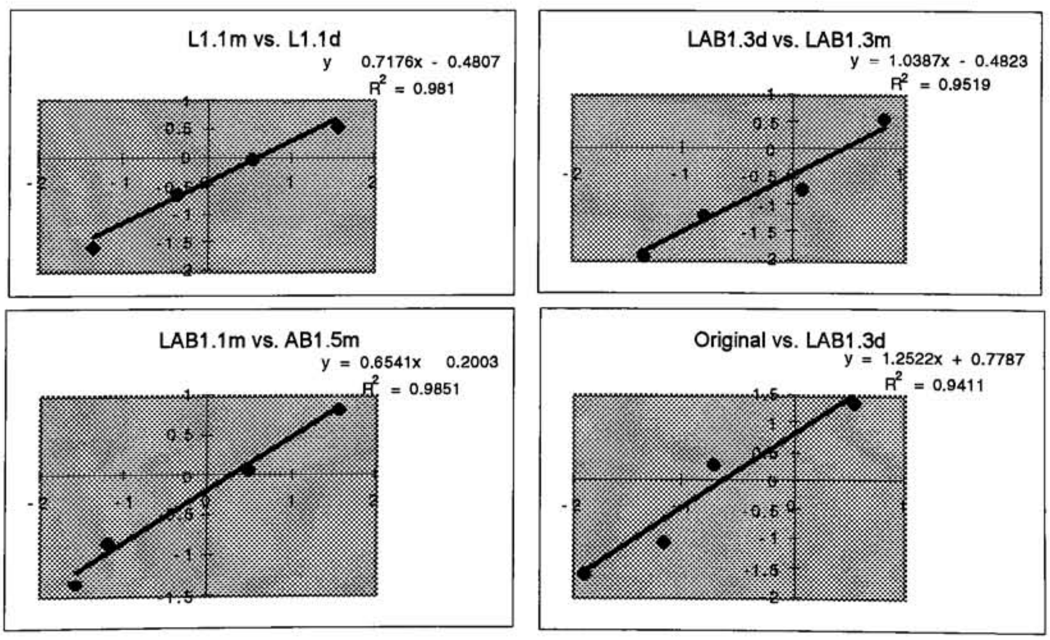


Figure 4-2-1.

If we plot the deviate values for each stimulus, as given in Table 4-2-4, as a function of the threshold values *Lc*, we obtain straight line regressions. Some plots are given in the next figure:

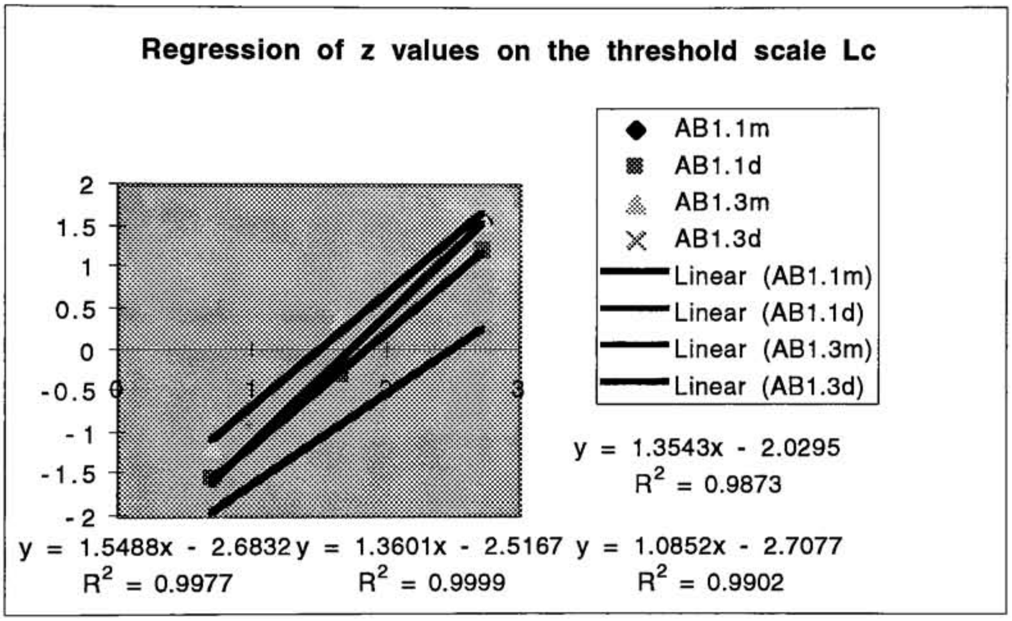


Figure 4-2-2.

The fact that each regression line is linear on the threshold scale is evidence of normality of the distribution for the stimuli.

Chapter 5.

Discussion and Results

After processing data, it was possible to formulate some conclusions.

The data was arranged such that the final classification was done in decreasing order of the median value corresponding to each image. The image with the smallest median value represents the most preferred image, between the set of 23 differently processed images. For a graphical representation of the preference, in order to have on the plots, the most preferred image appearing with the highest median value, the medians will be used with the sign changed. Also, to make possible a direct comparison between images, and to be able to find an average preference, the individual scales have been normalized, such that for all images the median values are normally distributed, with zero mean, and standard deviation equal with 1. These has been done using a linear transformation, that does not modifies the shape of the distributions.

Some differences have been noted, especially at the top of the hierarchy, between the two types of images, pictorial and graphics, and between the two stages of the experiment, when the original was or was not displayed for comparison.

For pictorial images, a tendency towards an increase in lightness and chromatic contrast has been noted, when the original was not shown on the CRT display. This can be seen in Figure 5-1, where the LAB1.1m is indicated as the most preferred image, and then AB1.1m .

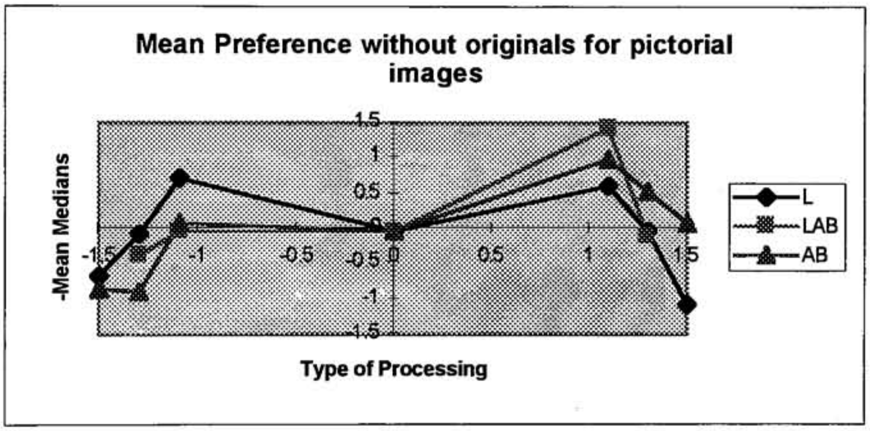


Figure 5-1.

Also, ***for all three pictorial images, the original,*** i.e., the image without having modified the lightness and chromatic contrast, ***was less preferred in the first stage of the experiment.*** In all three classifications it can be seen that the "Original" has jumped some positions towards the top, after the CRT image had been displayed.

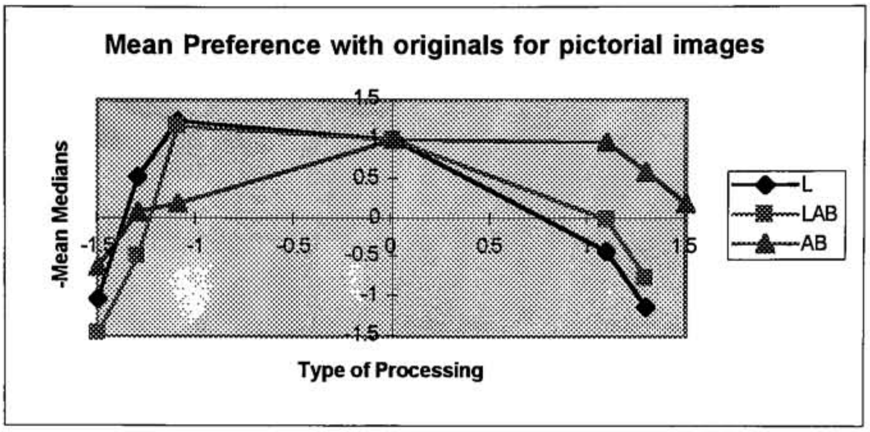


Figure 5-2.

For the images "Building" (Figure 5-3) and "Fruits" (Figure 5-4), the increase just in lightness contrast (L1.1m) was the most preferred, and the simultaneous increase in lightness and chromatic contrast was preferred for the image "Musicians" (Figure 5-5), when no originals where available. For this last image, "color memory" was probably the most used criterion in judging the quality of the image, and this might explain the lower position in hierarchy of the "Original", when no CRT image was used for comparison, and then the 11 positions jump, when the original was available. For the first two images, the "Original" is on position 6, and after seeing the original image on the display, in both cases, the new position is the third on top.

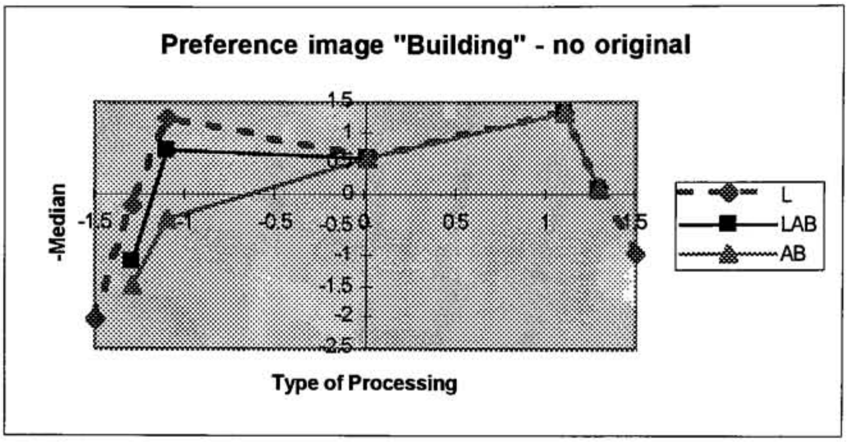


Figure 5-3.

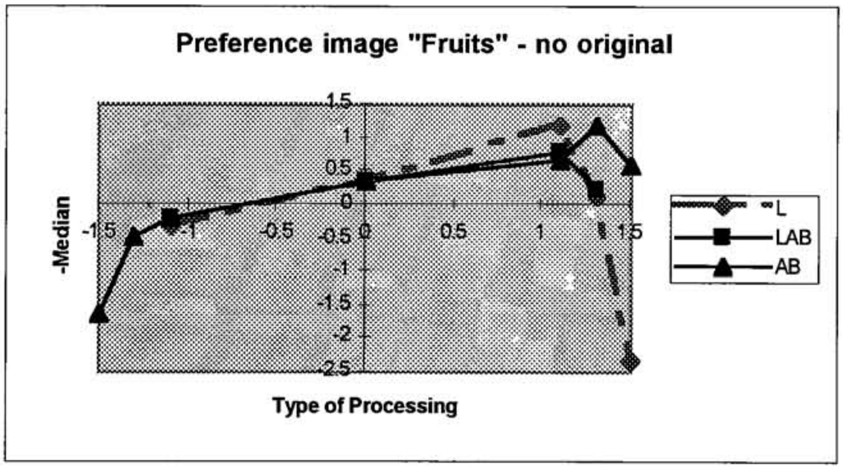


Figure 5-4.

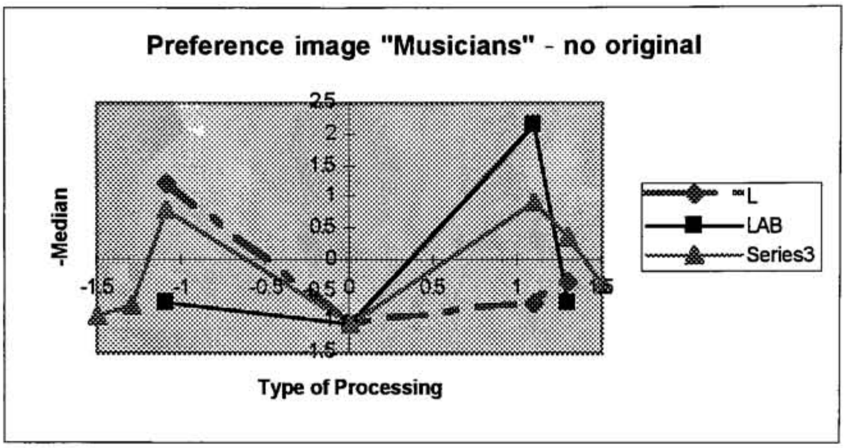


Figure 5-5.

For all three pictorial images, a decrease in lightness and chromatic contrast is preferred after the original CRT is seen.

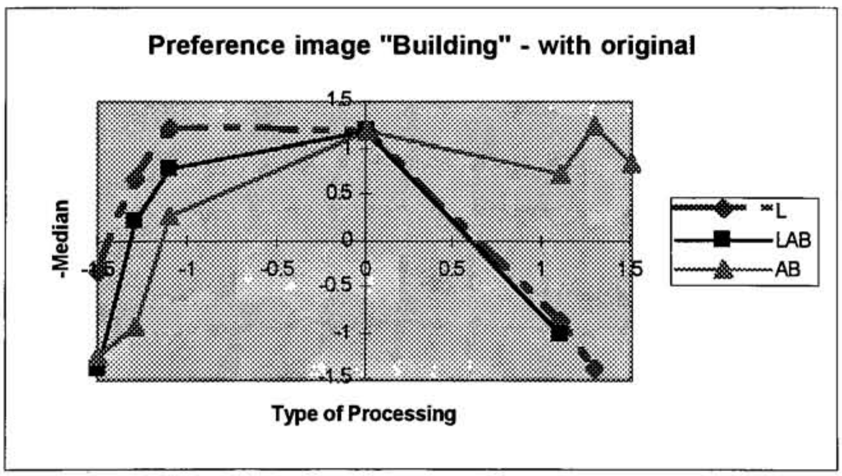


Figure 5-6.

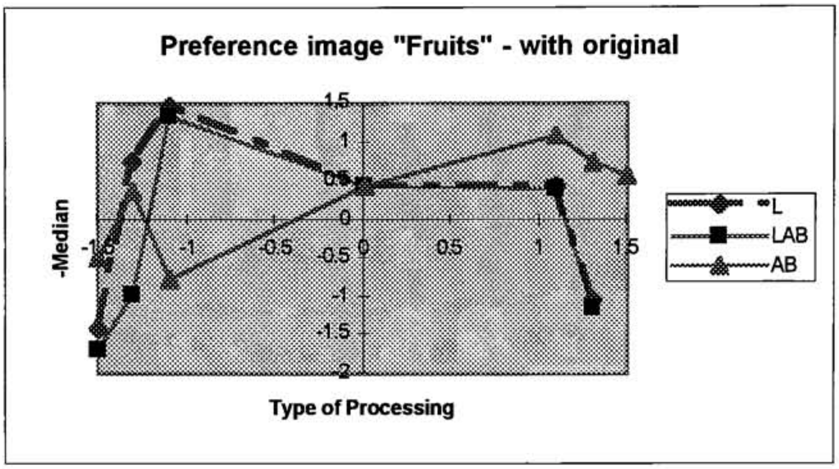


Figure 5-7.

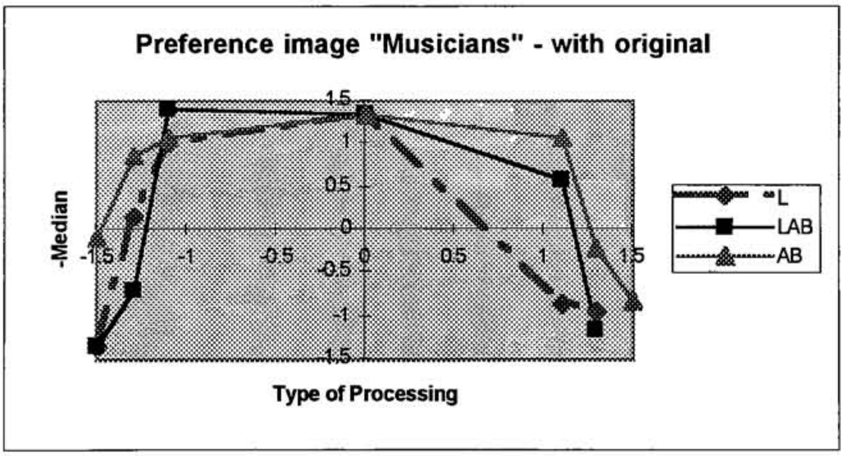


Figure 5-8.

At the lower end of the hierarchy, ***when the prints are seen without an original, the least preferred are those prints where the lightness and chromatic contrast was decreased excessively*** (AB1.5d, LAB1.5d, and L1.5d). This situation changes ***for the second stage of the experiment, the least preferred images being***

now those having the highest lightness and chromatic contrast (L1.5m and LAB1.5m).

For the graphics images, without having a CRT original was in favor for the simultaneous increase of the lightness and chromatic contrast, for both graphics images the LAB1.1m image being selected as the best one. In this stage of the experiment, when no original was available, the classification process was quite challenging for the observers, since no idea about how the original should look.

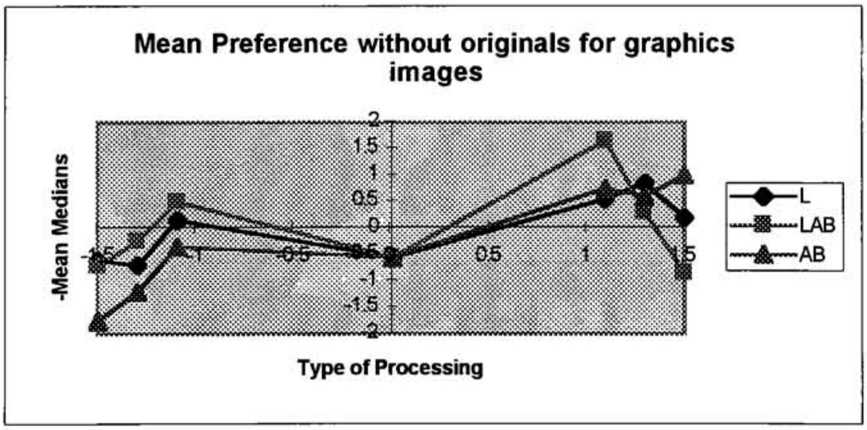


Figure 5-9.

When the CRT is made available, the preference moved towards less lightness and chromatic contrast, the LAB1.3d and AB1.5d being the most preferred combinations. For the graphics images with CRT original used for comparison, it can be observed that in the first ten positions are mostly those images for which the lightness and chromatic contrast was reduced (see Figure 5-11 and Figure 5-12). The same thing is depicted in Figure 5-10, where the mean preference, averaged across the graphics images is shown.

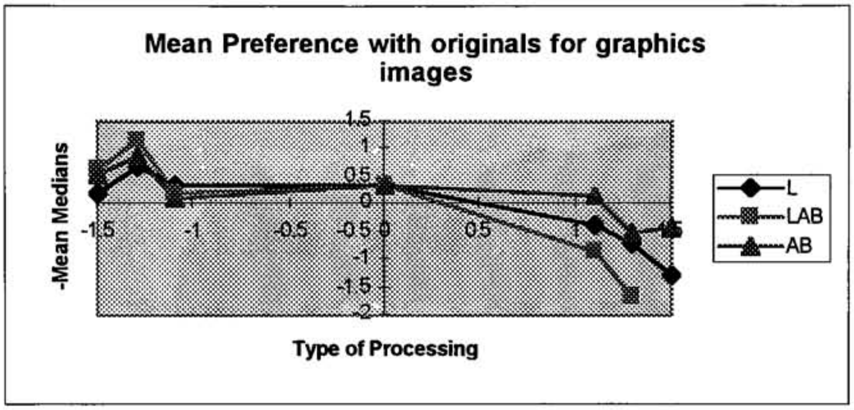


Figure 5-10.

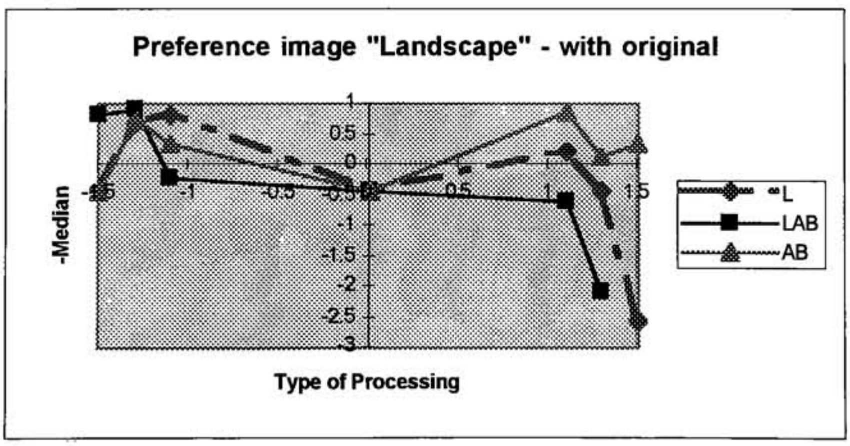


Figure 5-11.

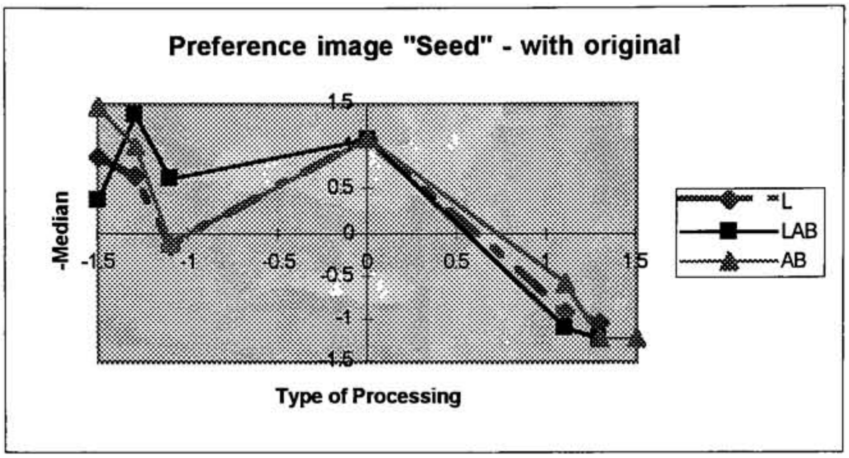


Figure 5-12.

A single exception, for the image "Landscape", where on the second place is the AB1.1m image, which means a slightly increased chroma is preferred. Otherwise, the next 5 positions for this image, indicate a decrease in lightness and chroma.

The viewing of the CRT image, makes the least preferred images, for both graphics images, those that have the highest lightness and chromatic contrast, LAB1.5m and L1.5m.

On the other hand, when the graphics images, "Landscape" and "Seed" are seen in isolation, on the last positions are those images having the lightness and chroma severely decreased; AB1.5d, LAB1.5d, and AB1.3d (Figure 5-13 and Figure 5-14).

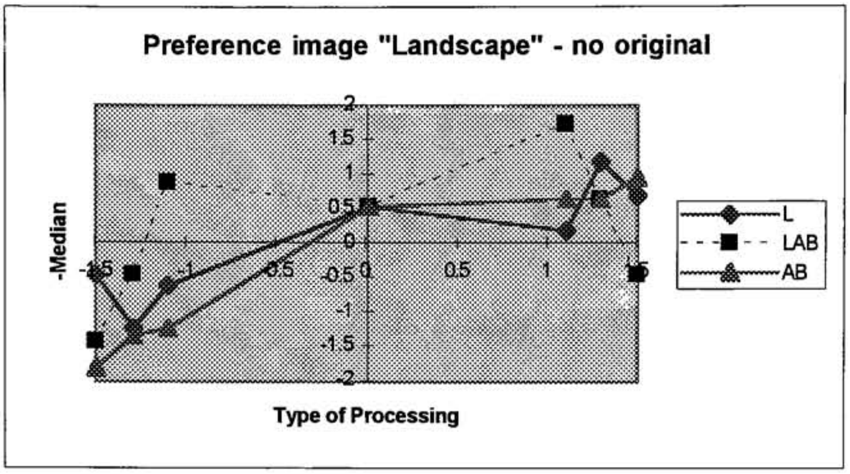


Figure 5-13.

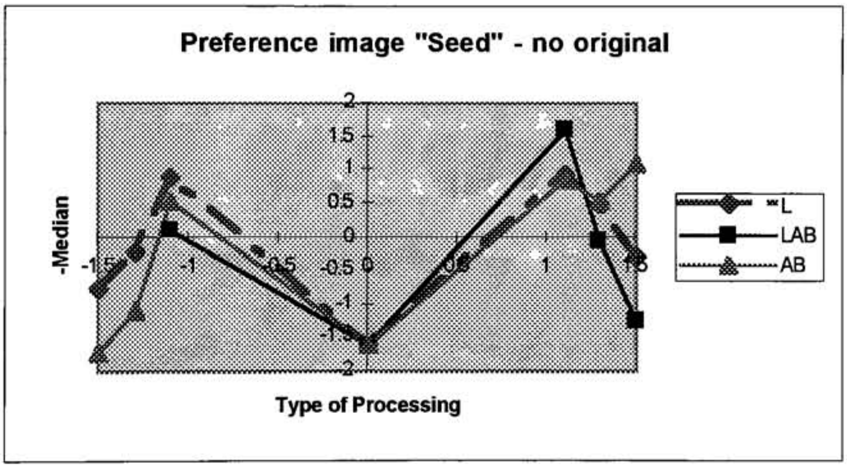


Figure 5-14.

Results

What can be concluded at the end of this analysis is that the observers' preference was dependent on the image content, some differences between pictorial images and the graphics ones, and also dependent on the type of the experiment, when the original was available for comparison or not. Actually, ***the highest differences in preference was between the two experiments***, as can be concluded if a comparison of the Figure5-9 vs. Figure 5-10, and Figure 5-1 vs. Figure 5-2 is performed.

For the graphics images without originals, a mean preference for the L1.1m processing has been noted, while for the same type of image, but with original available, the mean preference changed for LAB1.3d .

For the pictorial images, in the first phase of the experiment, when no original was used for comparison, the mean preference was LAB1.1m, and this mean preference changed into L1.1d and LAB1.1d, after the originals were used to compare.

These indicate that ***for the graphics images, without any a priori knowledge, people prefer more lightness contrast in the image, and for the pictorial images a simultaneous increase in chromatic and lightness contrast is preferred.***

An average preference across both types of images is represented in Figure 5-15 and Figure 5-16,

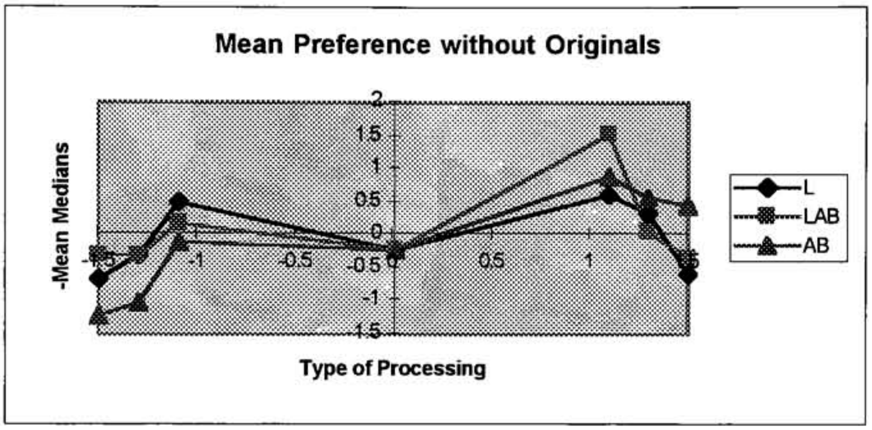


Figure 5-15.

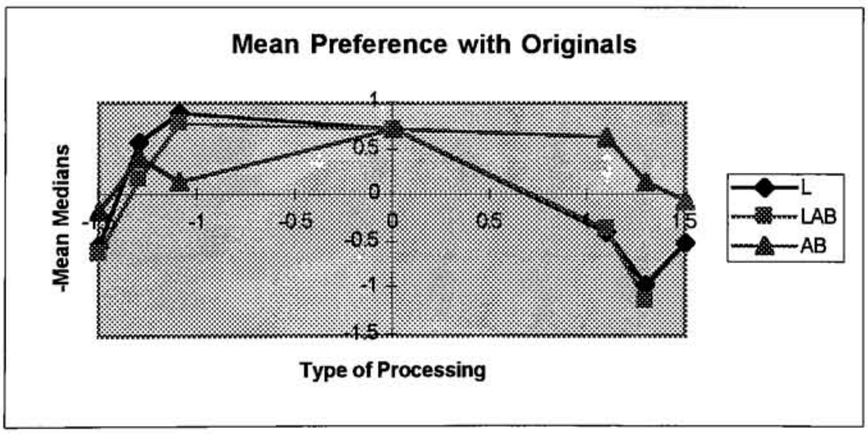


Figure 5-16.

It is perfectly clear that, ***the presence of the originals on CRT display is in favor of less lightness and chromatic contrast.***

Regarding the printer's performance, analyzing the Figure 24 from Appendix 3, that represents the MCDM, it can be concluded that the printer is quite stable; the min = 0.03, max = 2.21, median = 0.35, and mean = 0.49.

Also, in Appendix 3, the median values for ΔE_{94}, corresponding to all type of processing are given (see Figure 25 and Figure 26).

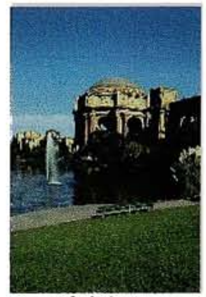
Appendix 1.

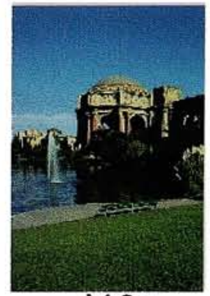
This appendix contains the images used in the experiment.

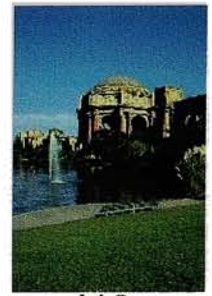

AB1.5d

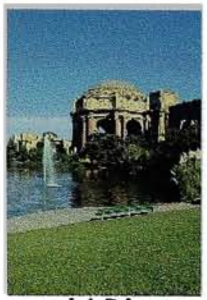

AB1.3d

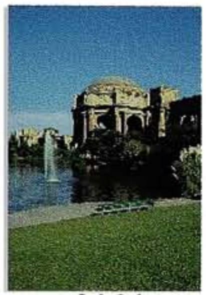

AB1.1d

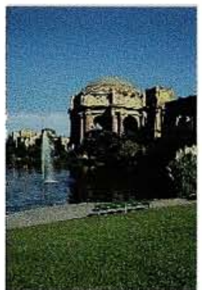

L1.1m

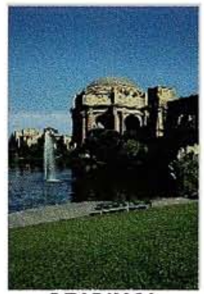

L1.3m

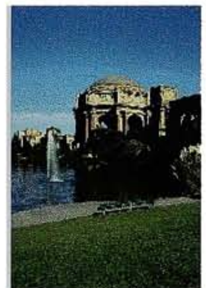

L1.5m

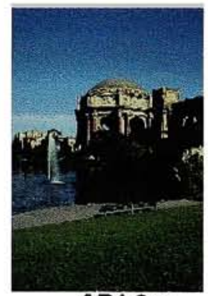

L1.5d

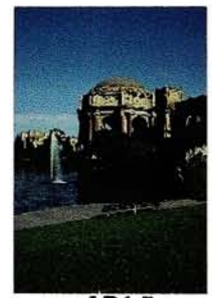

L1.3d

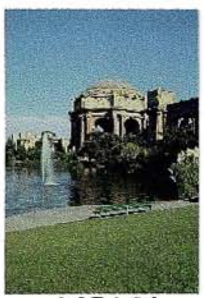

L1.1d


ORIGINAL

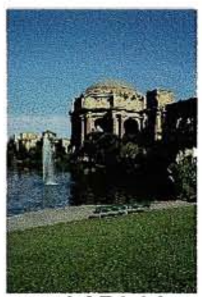

AB1.1m

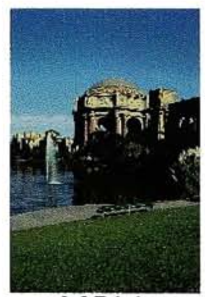

AB1.3m


AB1.5m


LAB1.5d


LAB1.3d


LAB1.1d


LAB1.1m


LAB1.3m


LAB1.5m


AB1.5d


L1.5d


LAB1.5d


AB1.3d


L1.3d


LAB1.3d


AB1.1d


L1.1d


LAB1.1d


ORIGINAL


AB1.1m


L1.1m


LAB1.1m


AB1.3m


L1.3m


LAB1.3m


AB1.5m


L1.5m


LAB1.5m


AB1.5d


L1.5d


LAB1.5d


AB1.3d


L1.3d


LAB1.3d


AB1.1d


L1.1d


LAB1.1d


ORIGINAL


AB1.1m


L1.1m


LAB1.1m


AB1.3m


L1.3m


LAB1.3m


AB1.5m


L1.5m


LAB1.5m


AB1.5d


L1.5d


LAB1.5d


AB1.3d


L1.3d


LAB1.3d


AB1.1d


L1.1d


LAB1.1d


ORIGINAL


AB1.1m

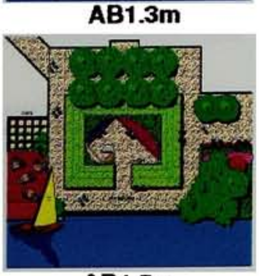

L1.1m


LAB1.1m


AB1.3m


L1.3m


LAB1.3m


AB1.5m


L1.5m


LAB1.5m


AB1.5d


AB1.3d


AB1.1d


AB1.1m


AB1.3m


AB1.5m


L1.5d


L1.3d


L1.1d


ORIGINAL


L1.1m


L1.3m


L1.5m


LAB1.5d


LAB1.3d


LAB1.1d


LAB1.1m


LAB1.3m


LAB1.5m

Appendix 2.

For the colorimetric calculations, necessary to characterize the printer's behavior, an average of five prints representing a 5x5x5 target was used. These five prints represent the target without any a priori processing.

The flowchart of the calculations can be seen in the following figures:

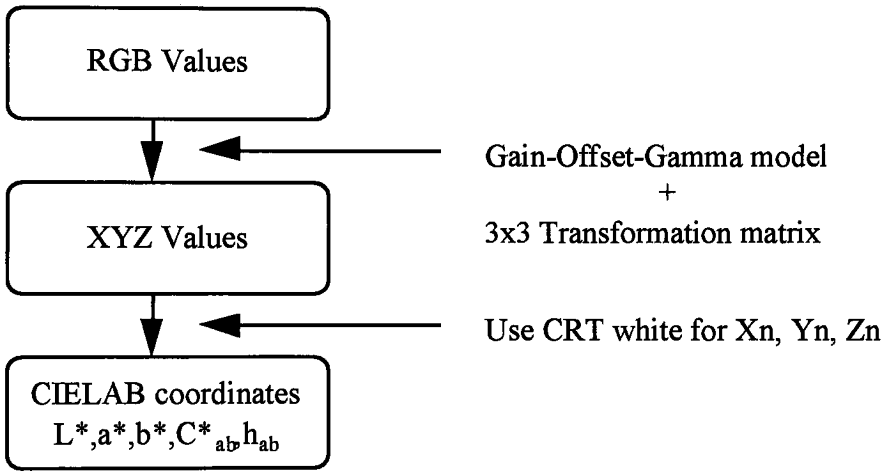


Figure 1.The flowchart of the calculations for CRT.

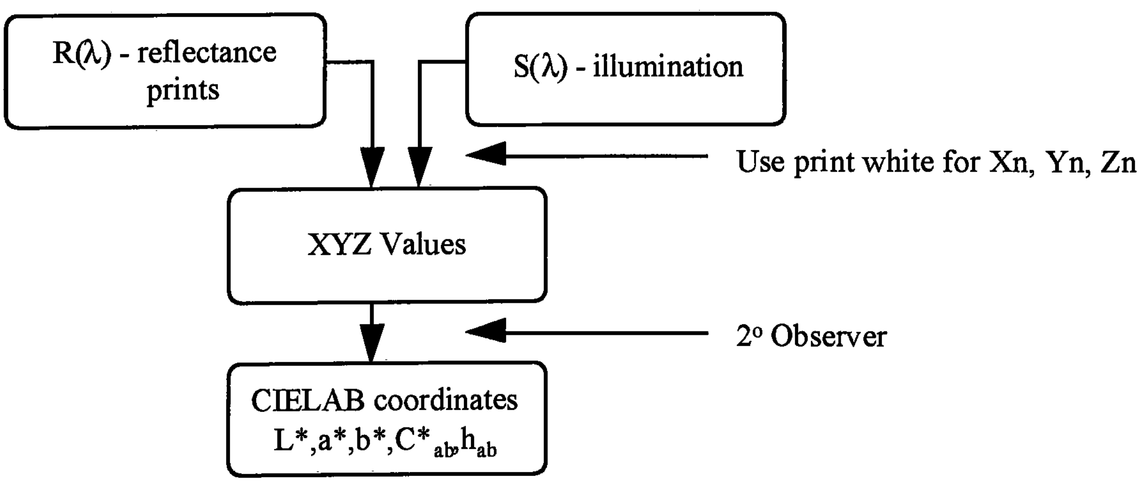


Figure 2. The flowchart of the calculations for prints.

The spectral power distribution of the source, and the reflectances of the prints, were considered for the wavelength interval, between 380 and 730 nm, with 4nm increments. The spectral power distribution of the source represents the average of 20 measurements performed using a Photo Research PR-650 spectrophotometer, and the reflectances of the prints were measured using GreTag.

The next figures will try to give us a better idea about the relationship between the CIELAB coordinates of the CRT image and the corresponding printed images of the 5x5x5 target. An increase in L* values for black patches, and the decrease in L* values for the yellow ones can easily be noted, in Figure 3 and Figure 4.

The next figures, starting with Figure 5 represent the mapping of the other CIELAB coordinates from CRT onto the prints. The differences Δa^*, Δb^*, ΔL^*, ΔC^*_{ab}, and Δh_{ab} represent the differences between the corresponding values for the MeanOriginal and CRT.

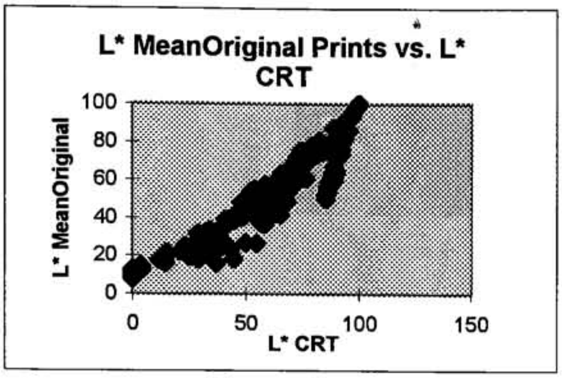


Figure 3.

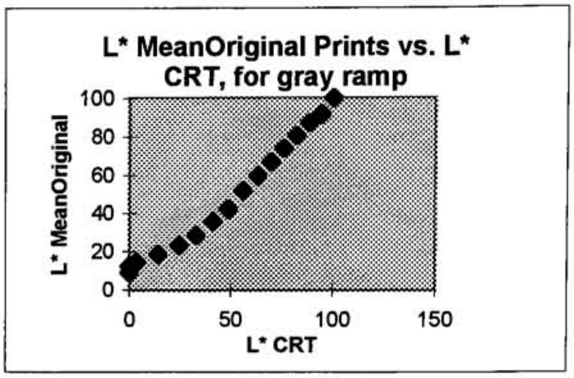


Figure 4.

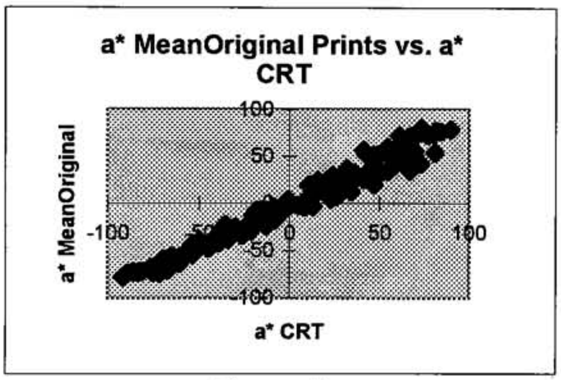


Figure 5.

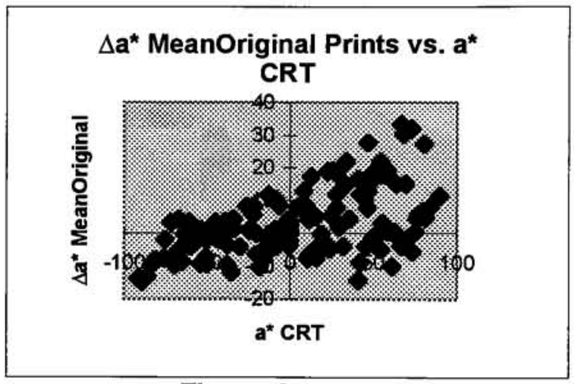


Figure 6.

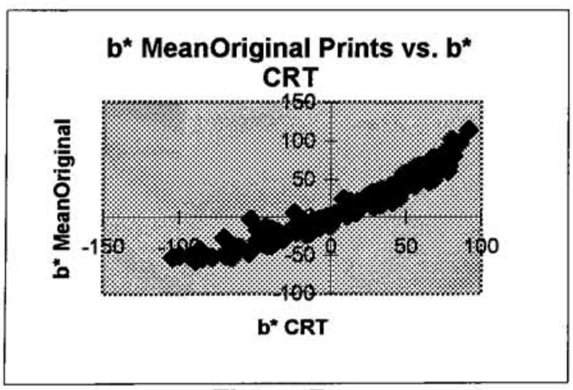


Figure 7.

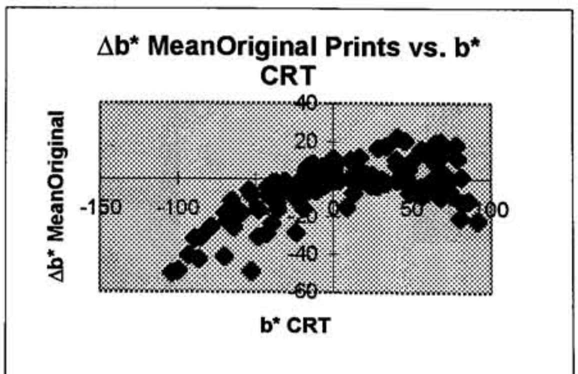


Figure 8.

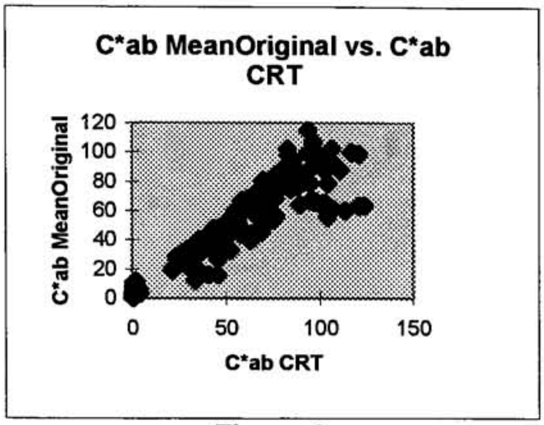


Figure 9.

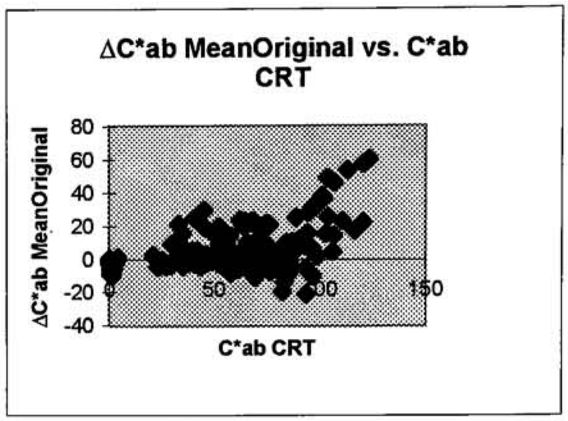


Figure 10.

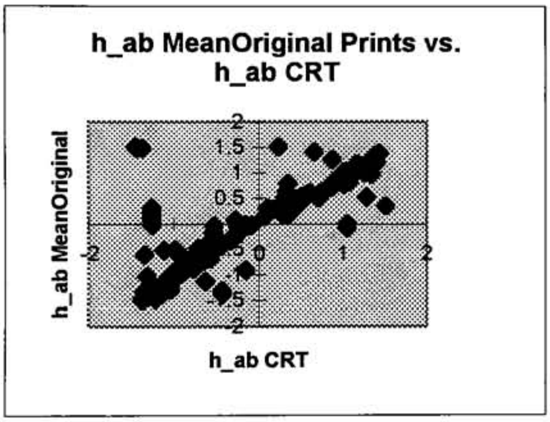


Figure 11.

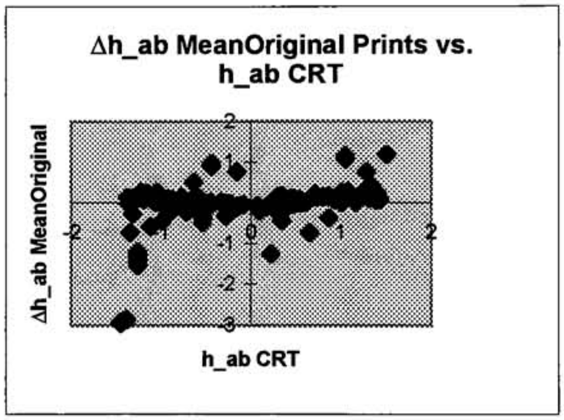
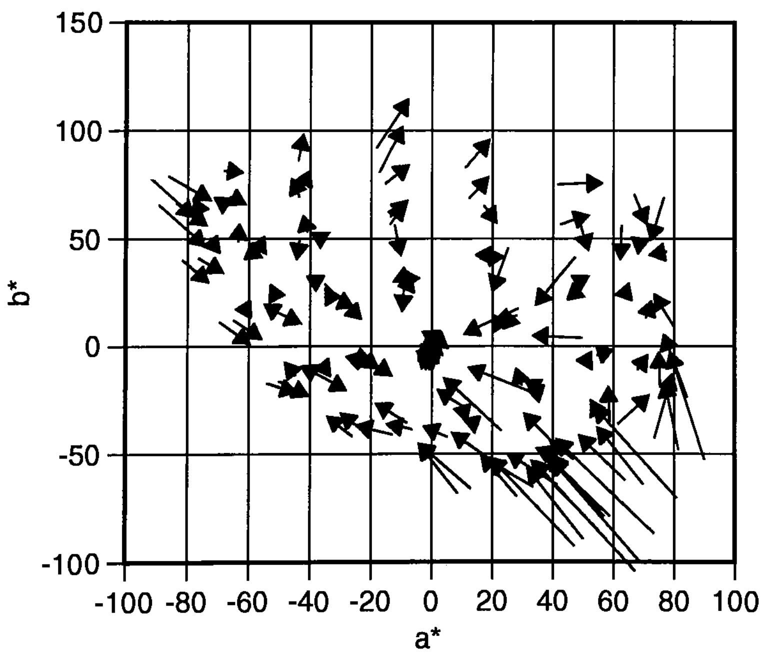


Figure 12.

Figures 13 (a) and (b) represent the shift in chroma and the relationship between the (C*ab, L*) coordinates of the CRT and the corresponding (C*ab, L*) mean values for the print. The arrows point from the corresponding CRT values towards the print values.

As can be seen in the next figure, the highest shift in chroma can be noted from red-blue towards green-yellow, in the right bottom part of the figure. Also, in the left upper part, from yellow-green towards blue-red.

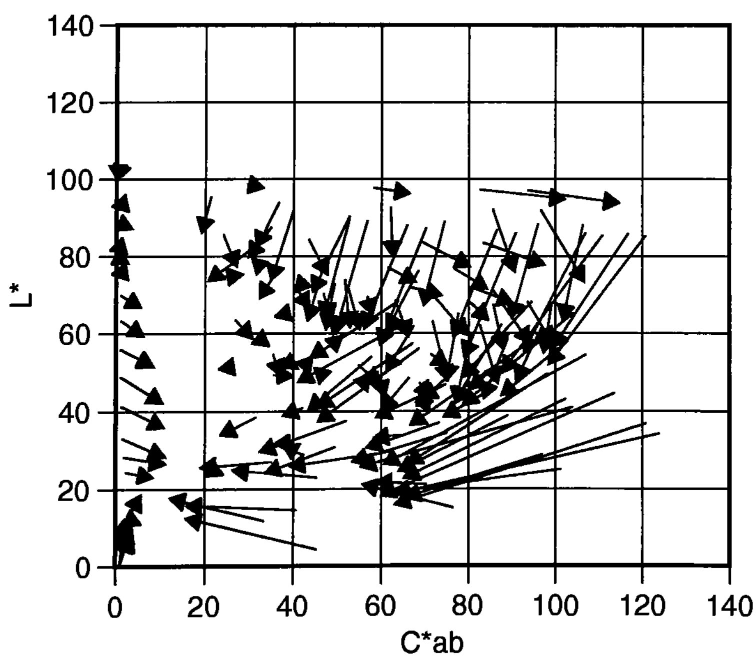
Figure 13 (a). Error vectors for (a*, b*).



For the neutrals, an increase in lightness can be observed for low values , then a decrease in lightness in the [25, 70] range. For the highest values of the lightness, an increase of the C^*_{ab} can be noted, but generally the corresponding C^*_{ab} values of the prints are much lower as compared with C^*_{ab} values of the CRT. The highest differences in C*ab can be seen for the lightness values in the [20, 40] range.

Also, for the C^*_{ab} in the [80, 120] range, the highest variation in lightness is indicated in the figure.

Figure 13 (b). Error vectors for (C*ab, L*).



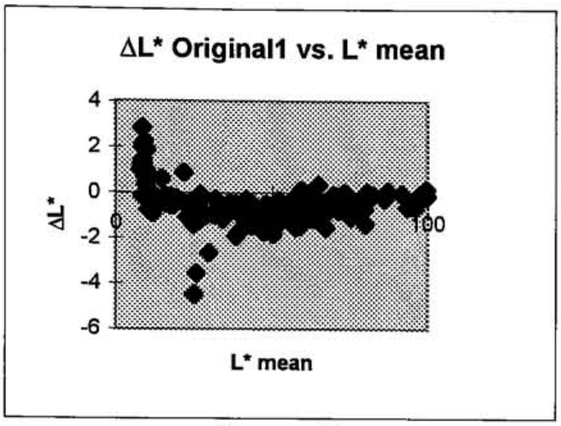


Figure 14.

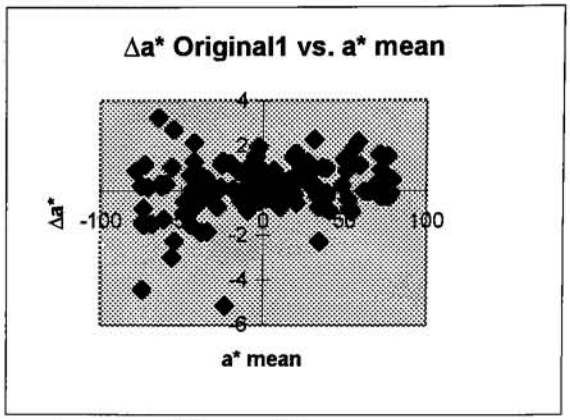


Figure 15.

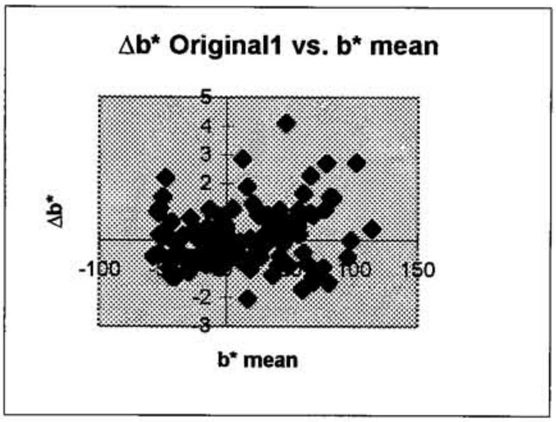


Figure 16.

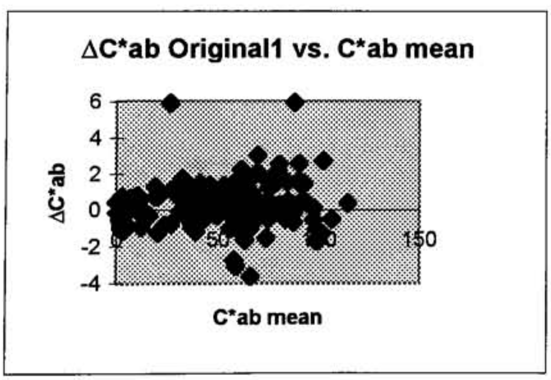


Figure 17.

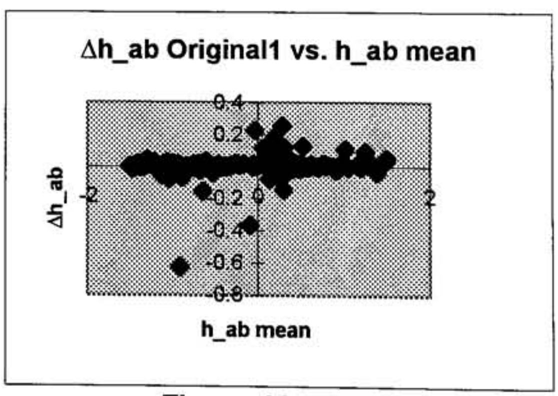


Figure 18.

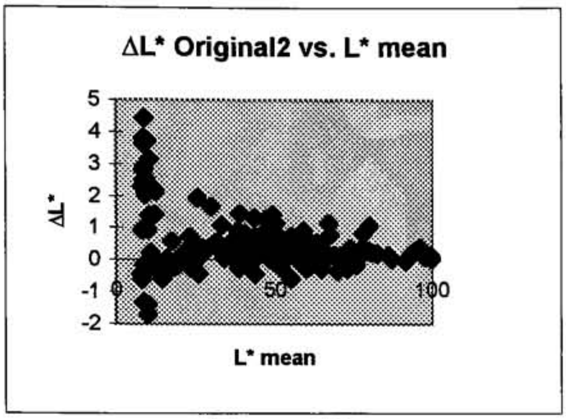


Figure 19.

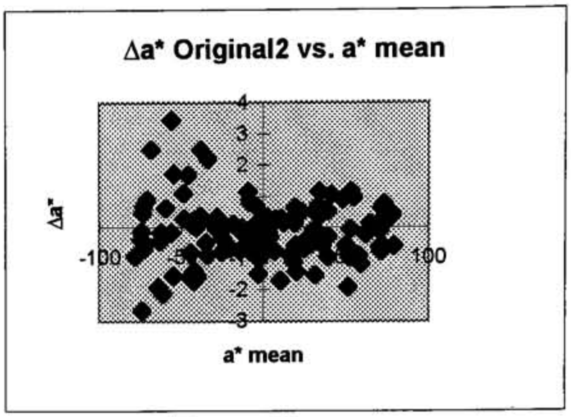


Figure 20.

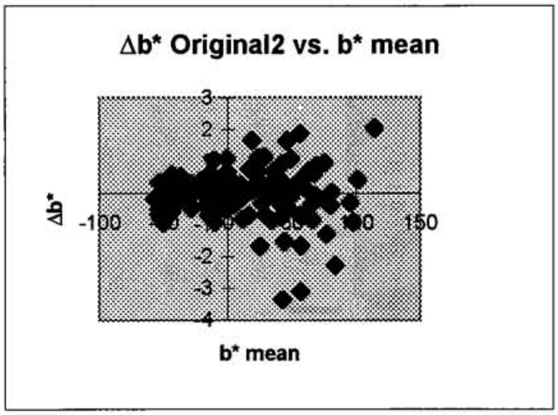


Figure 21.

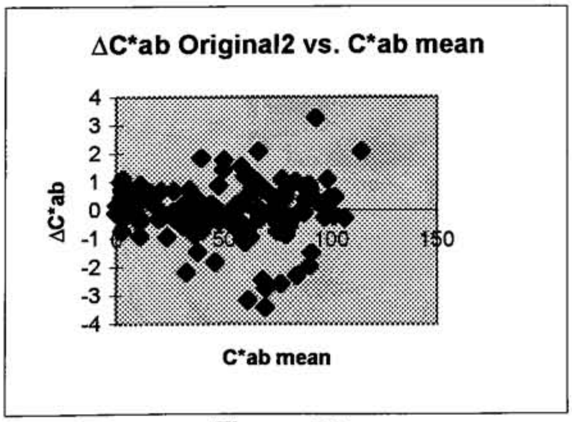


Figure 22.

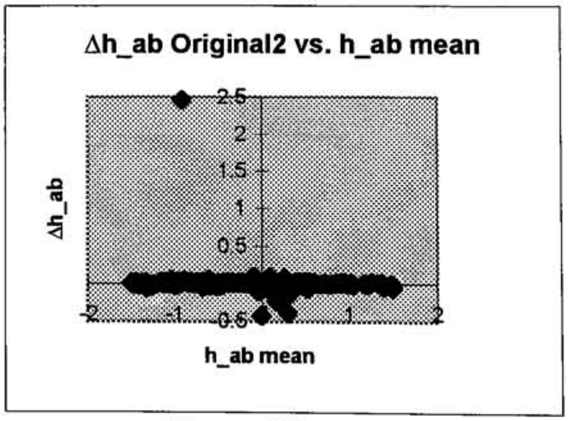


Figure 23.

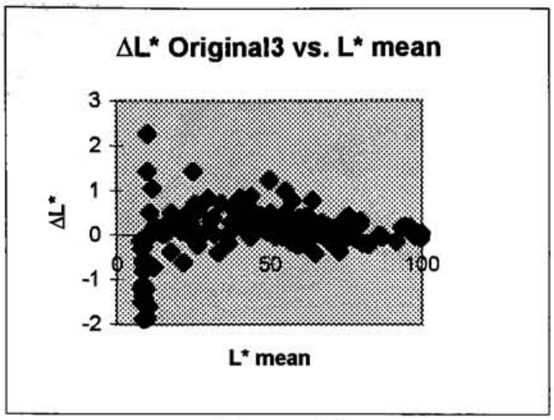


Figure 24.

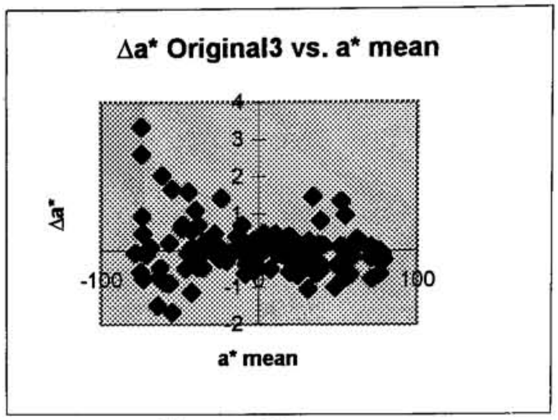


Figure 25.

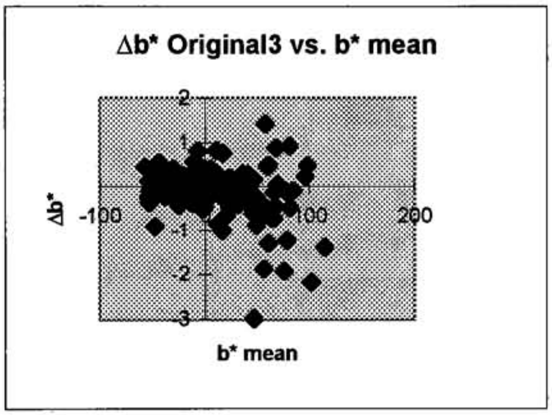


Figure 26.

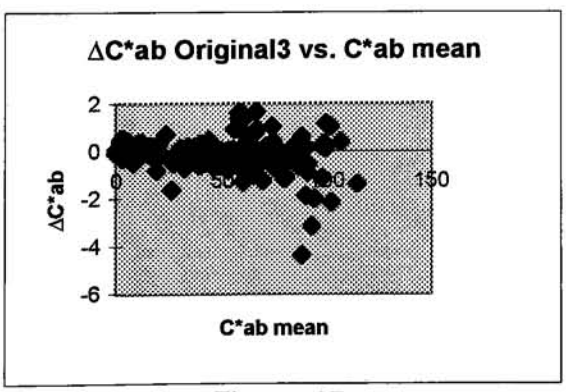


Figure 27.

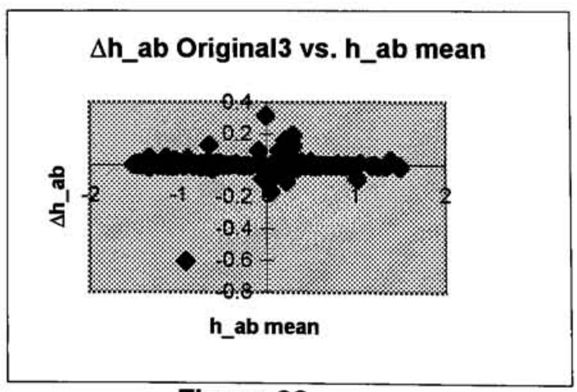


Figure 28.

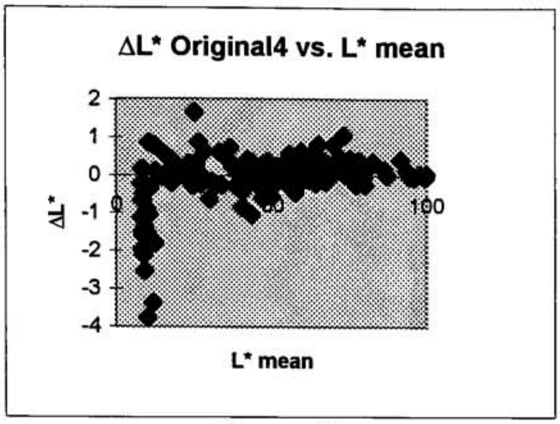


Figure 29.

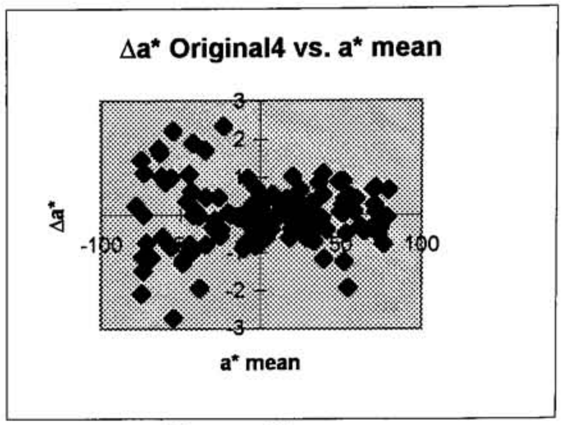


Figure 30.

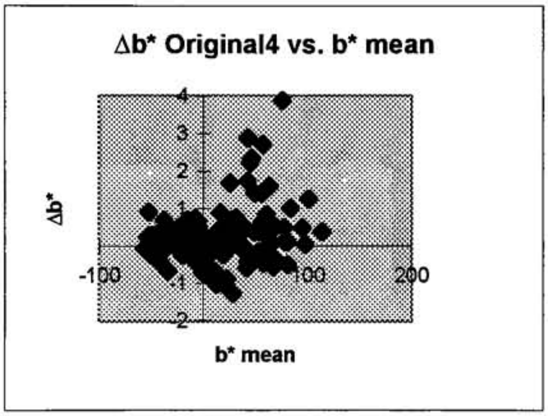


Figure 31.

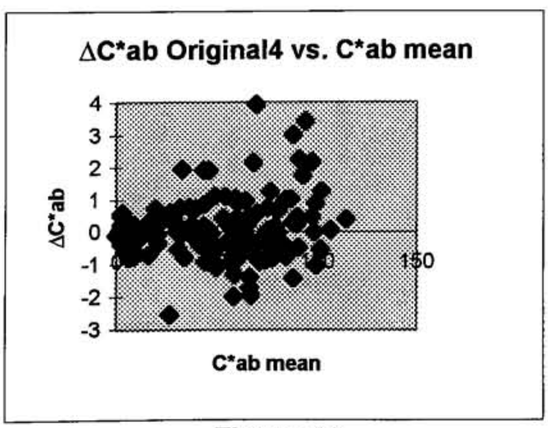


Figure 32.

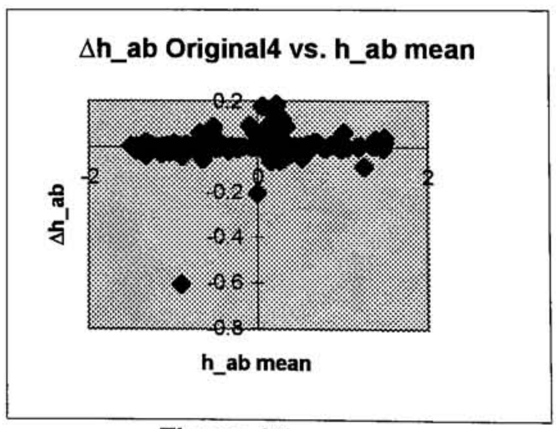


Figure 33.

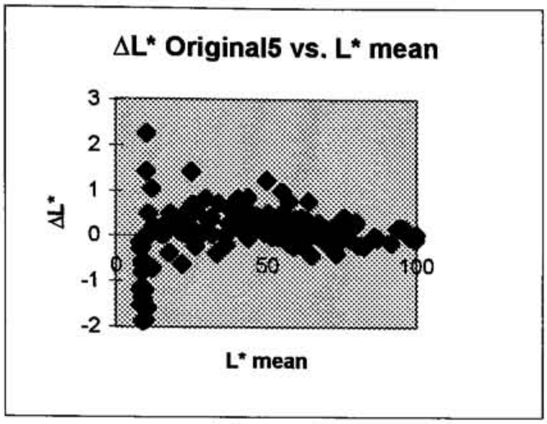


Figure 34.

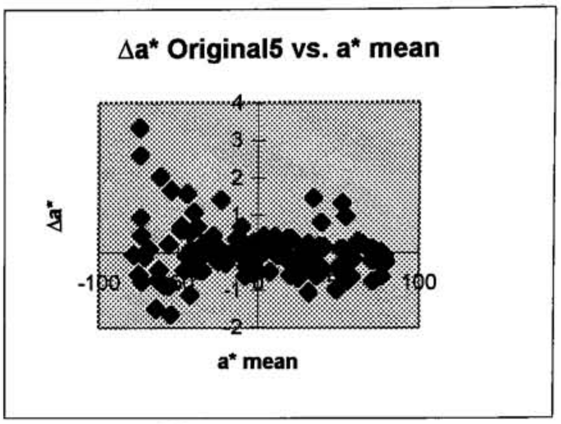


Figure 35.

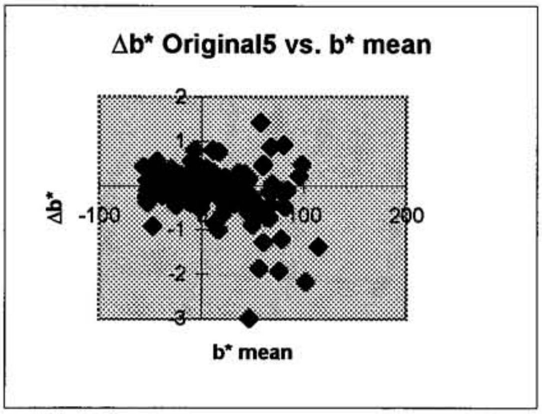


Figure 36.

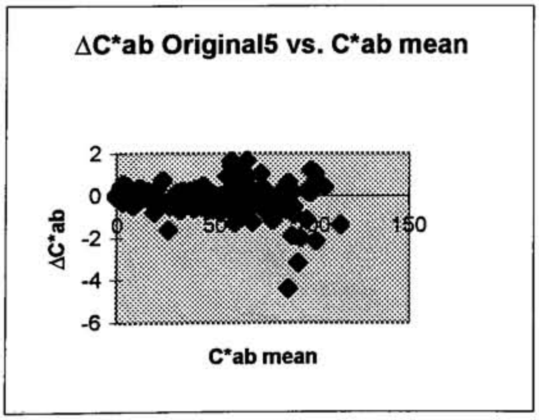


Figure 37.

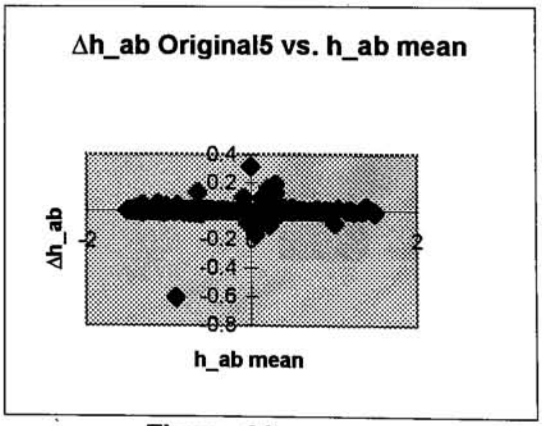


Figure 38.

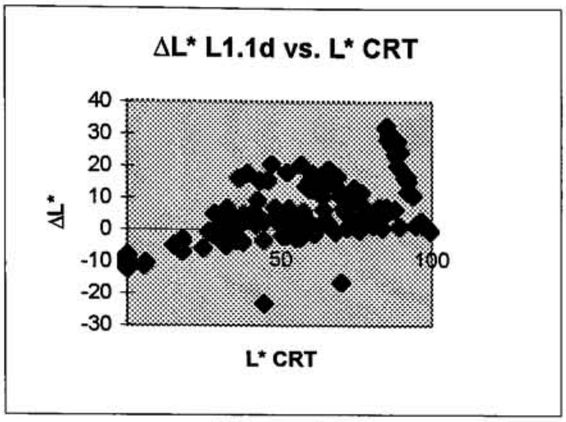


Figure 39.

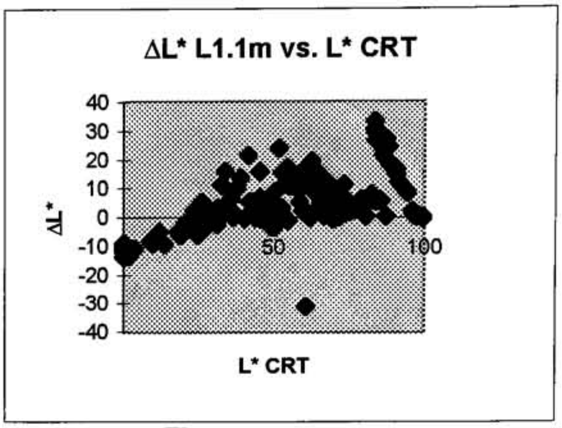


Figure 40.

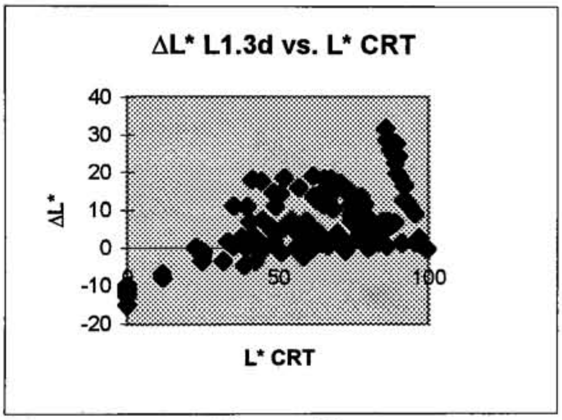


Figure 41.

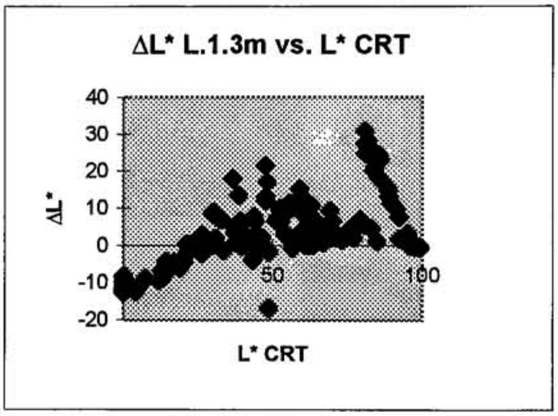


Figure 42.

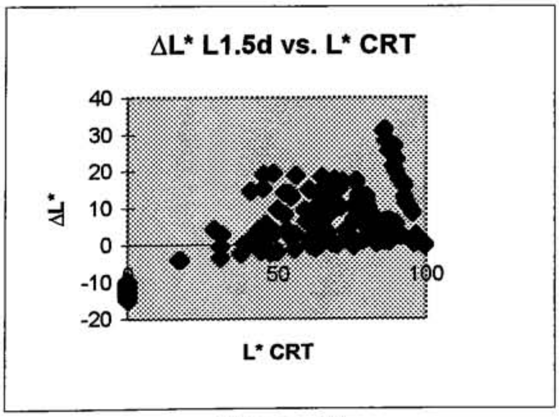


Figure 43.

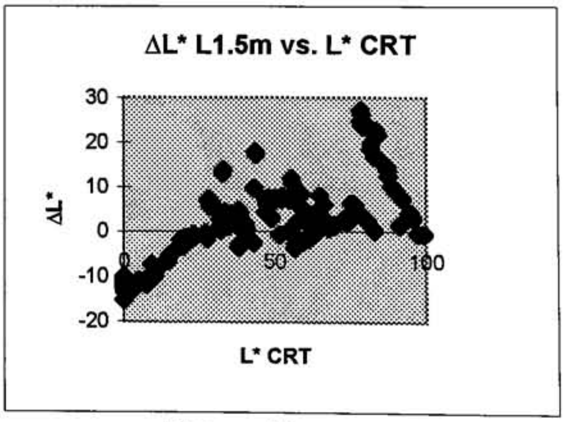


Figure 44.

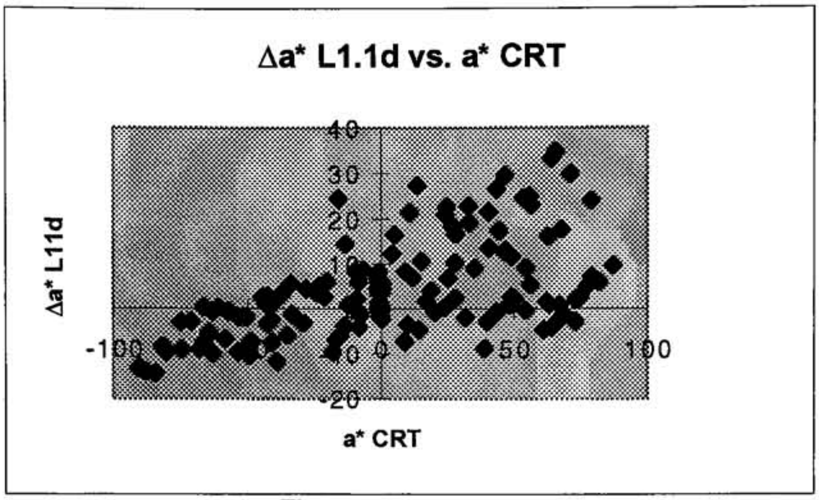


Figure 45.

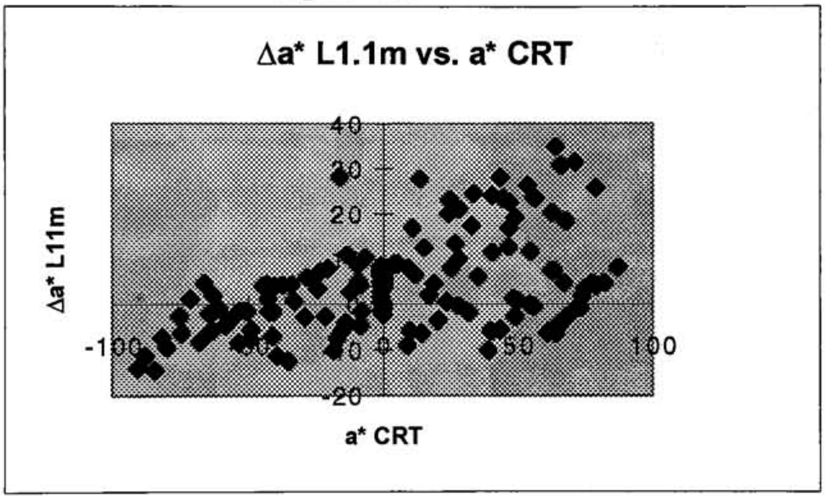


Figure 46.

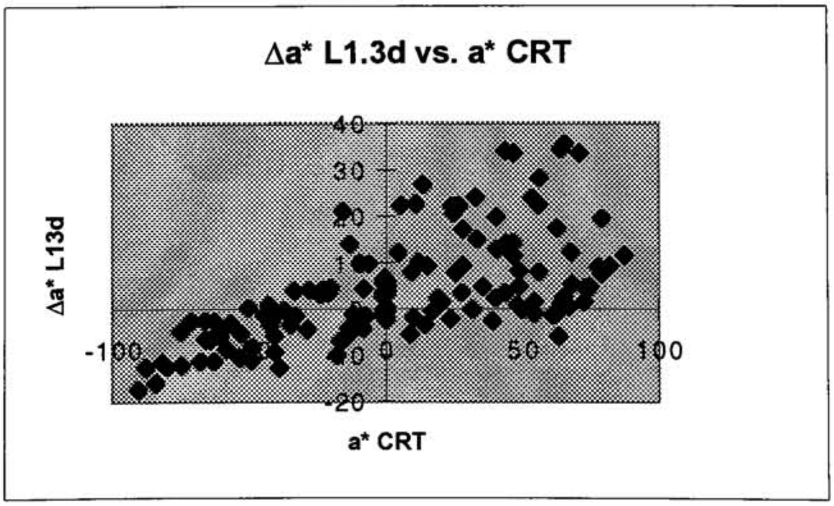


Figure 47.

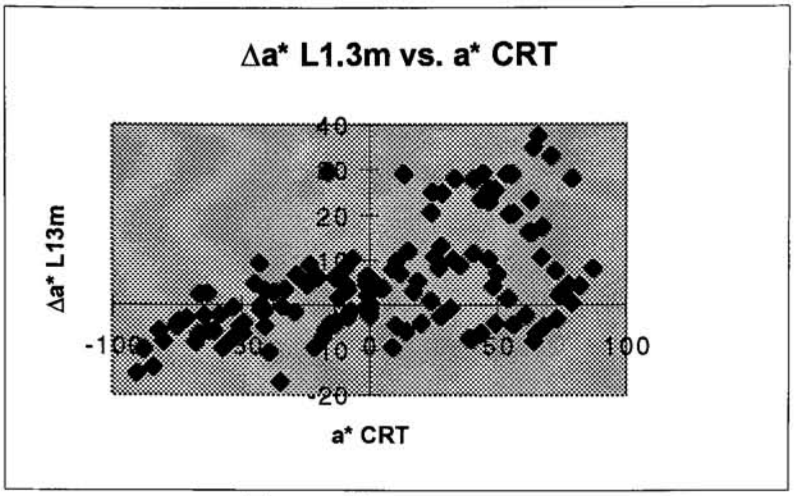


Figure 48.

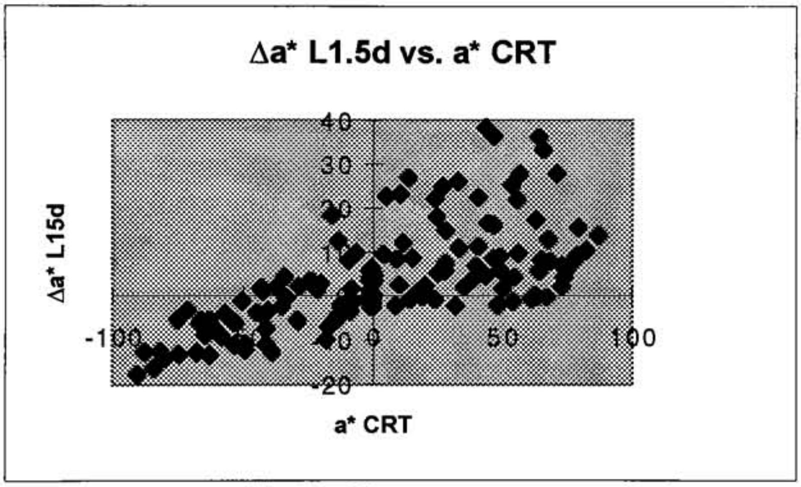


Figure 49.

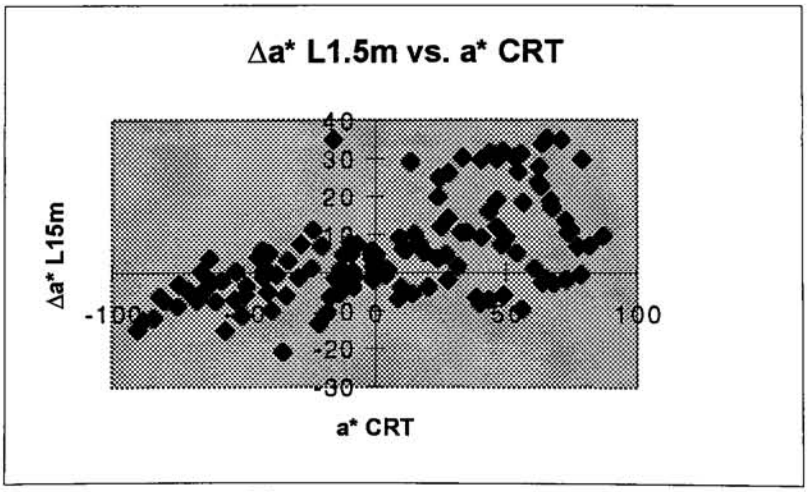


Figure 50.

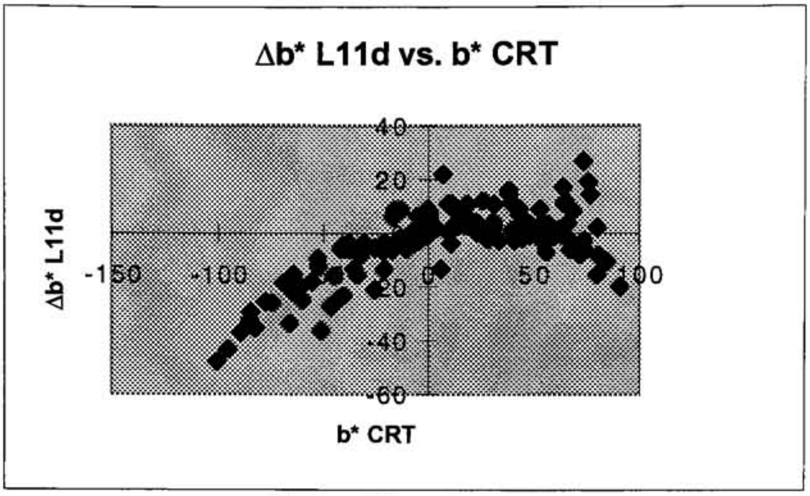


Figure 51.

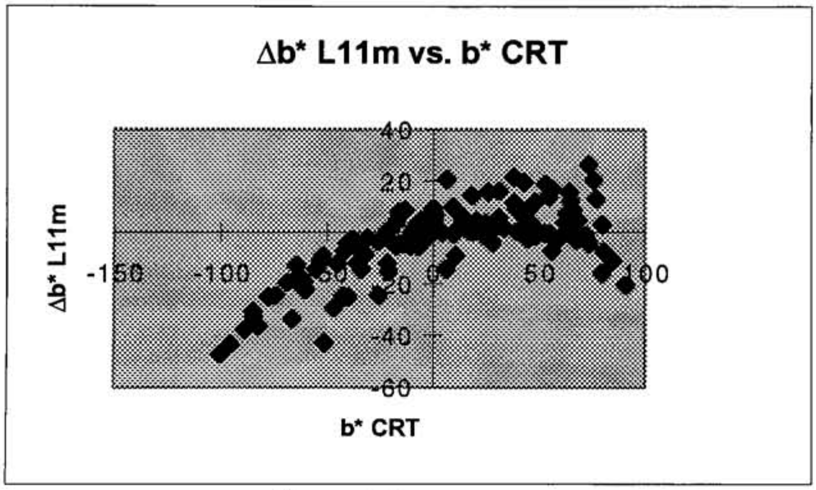


Figure 52.

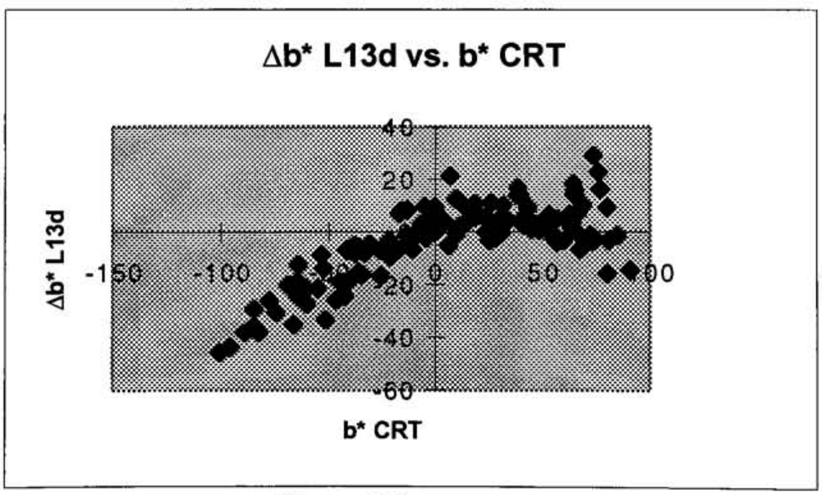


Figure 53.

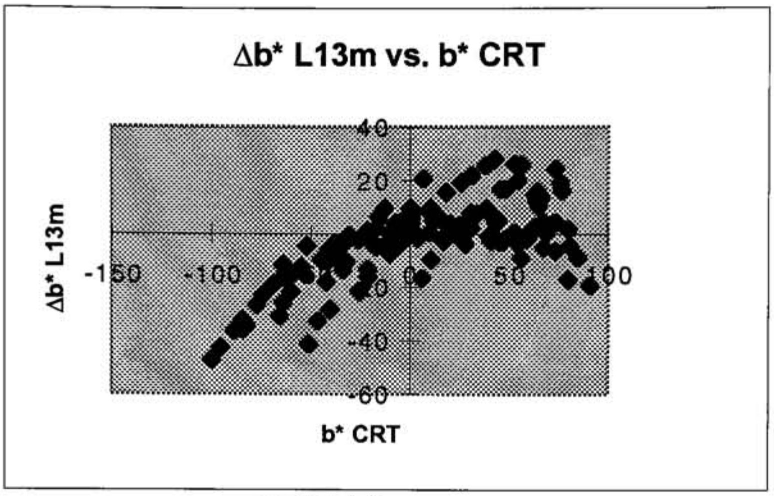


Figure 54.

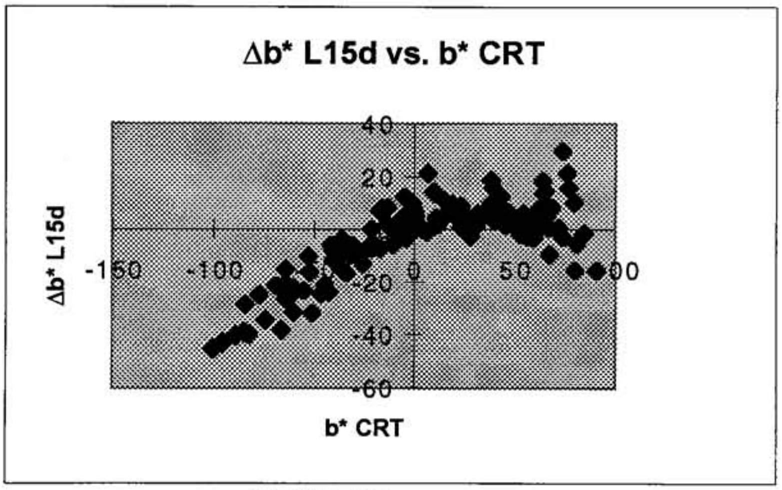


Figure 55.

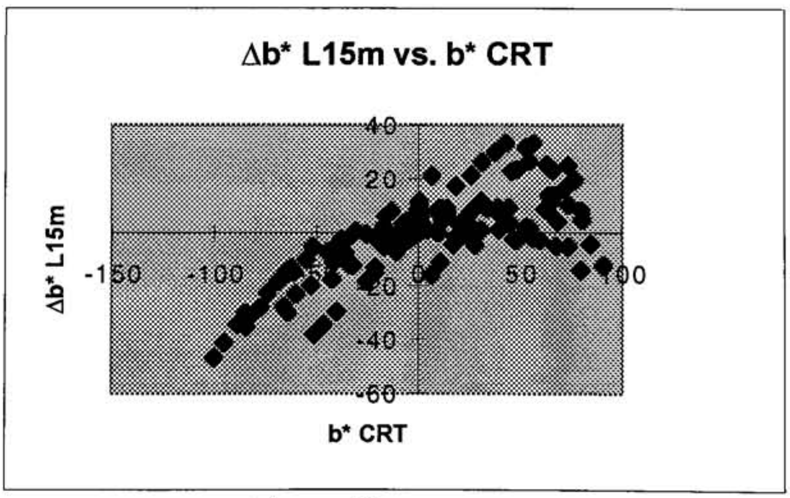


Figure 56.

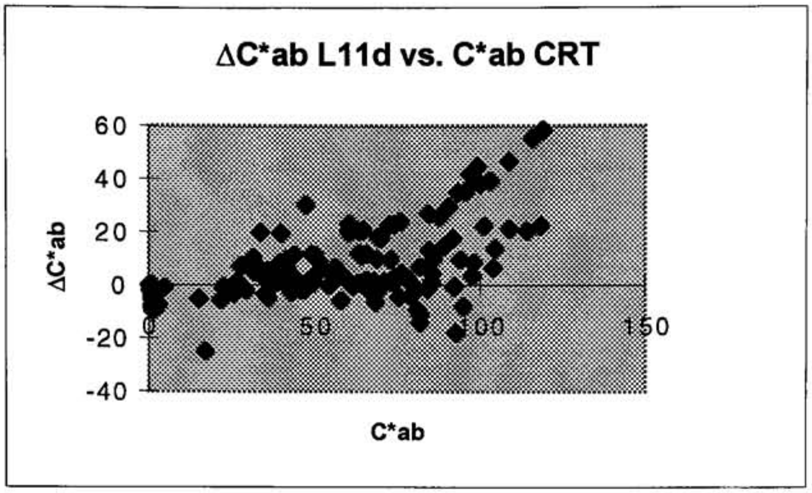


Figure 57.

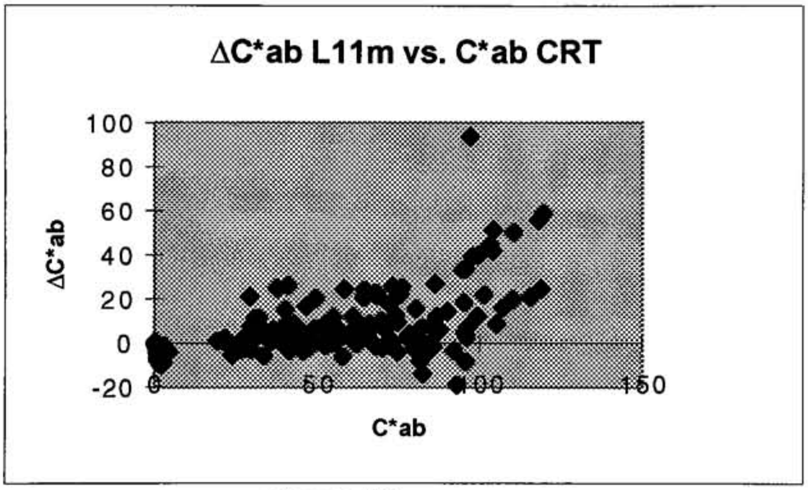


Figure 58.

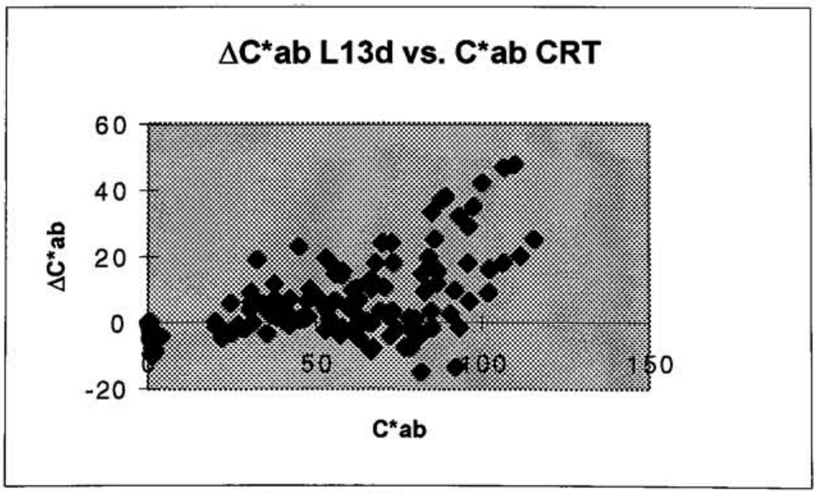


Figure 59.

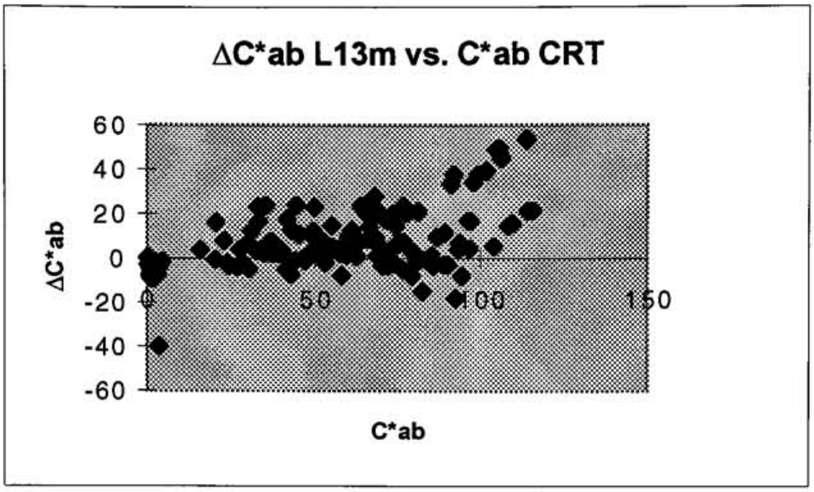


Figure 60.

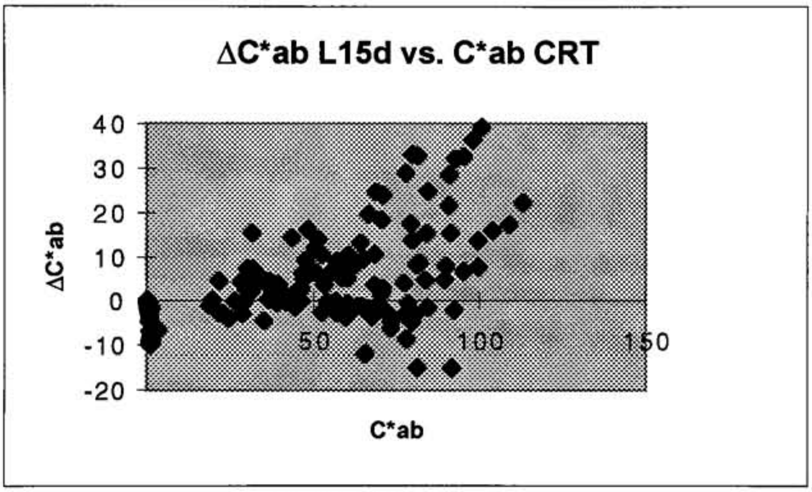


Figure 61.

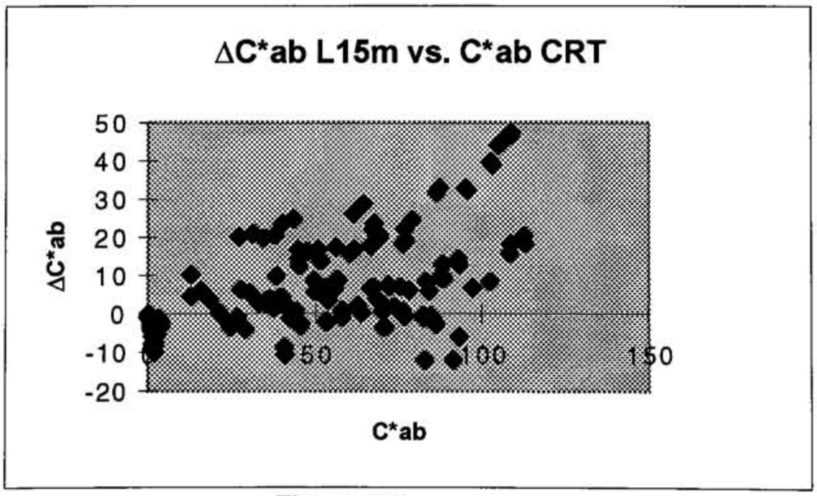


Figure 62.

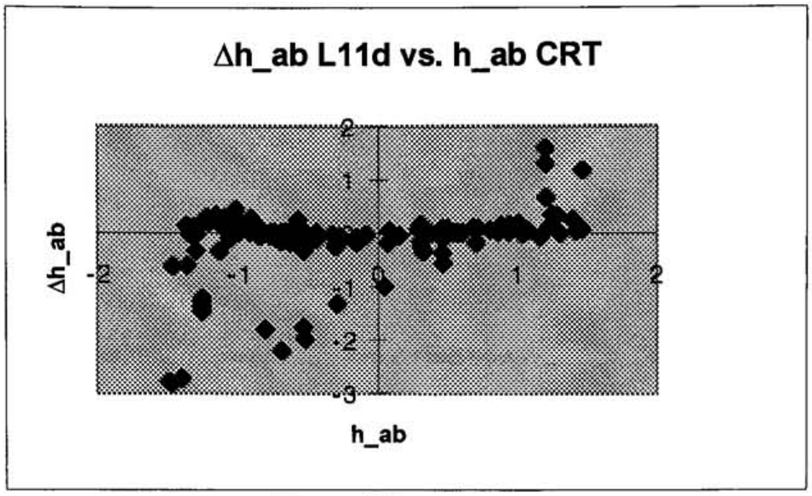


Figure 63.

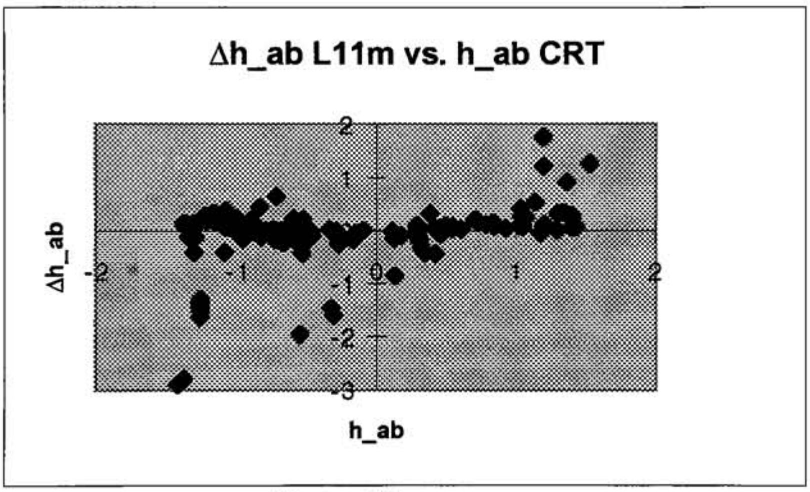


Figure 64.

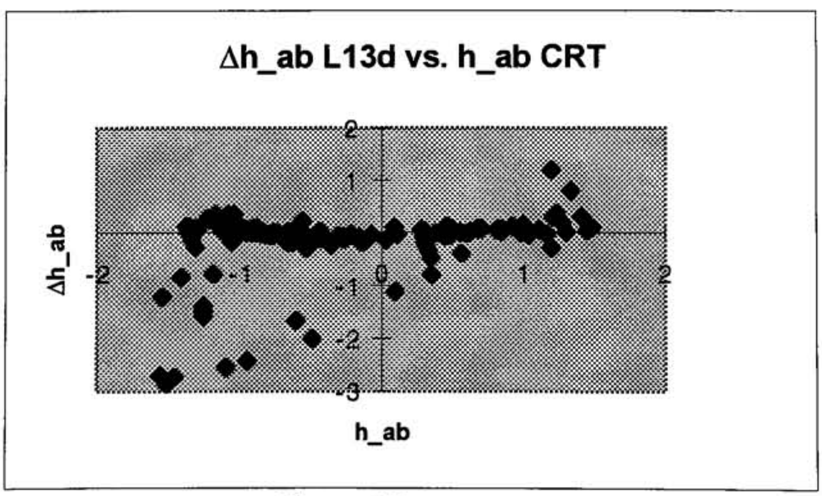


Figure 65.

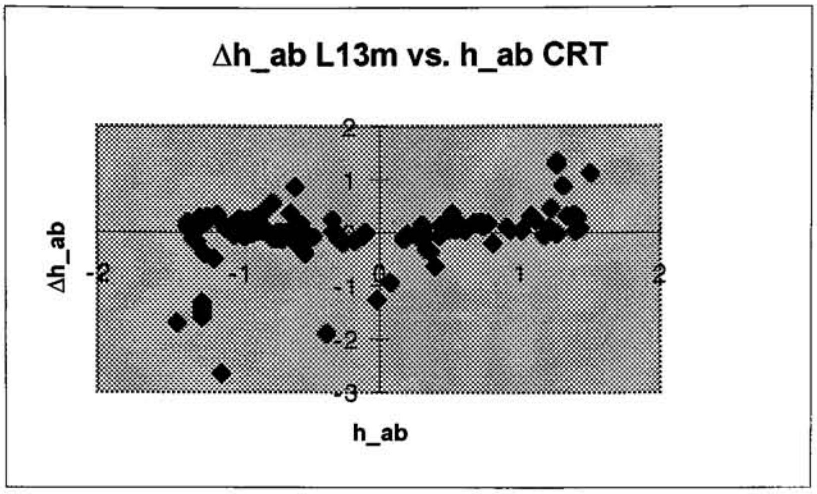


Figure 66.

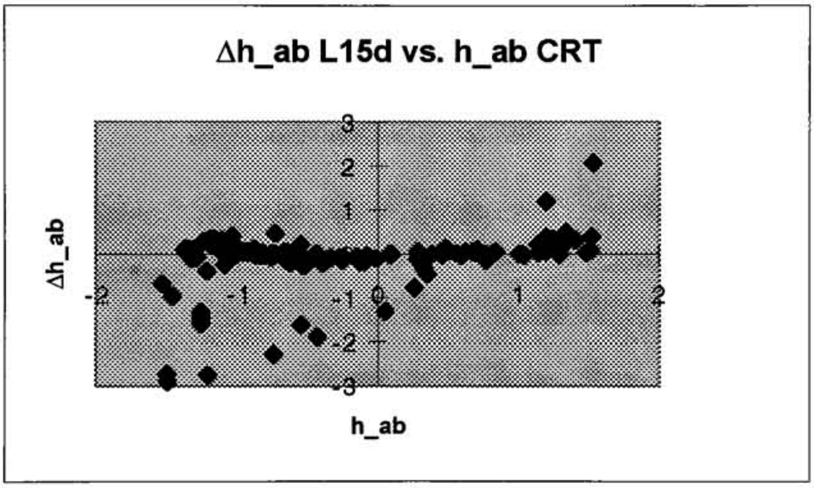


Figure 67.

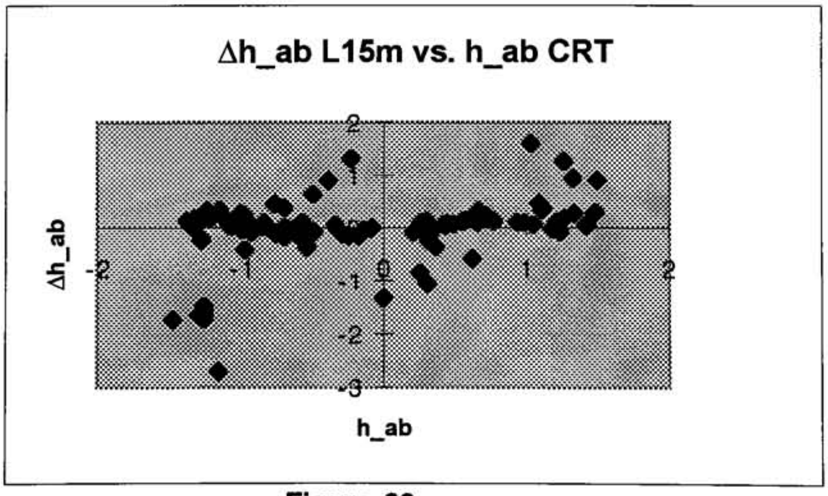


Figure 68.

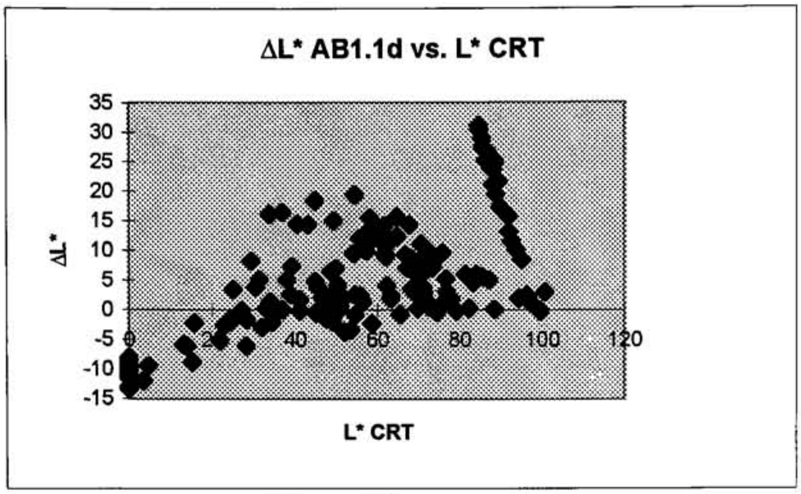


Figure 69.

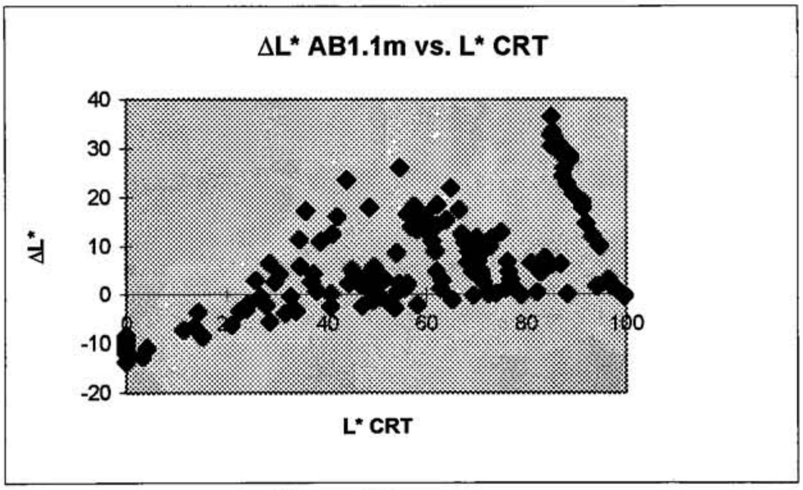


Figure 70.

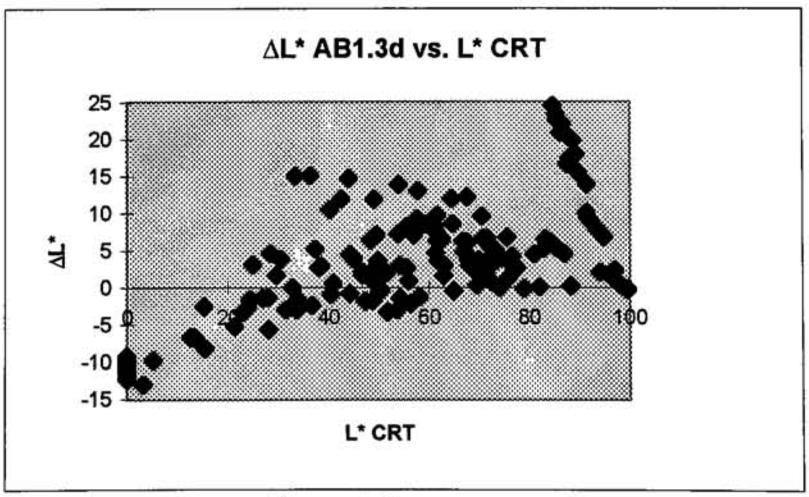


Figure 71.

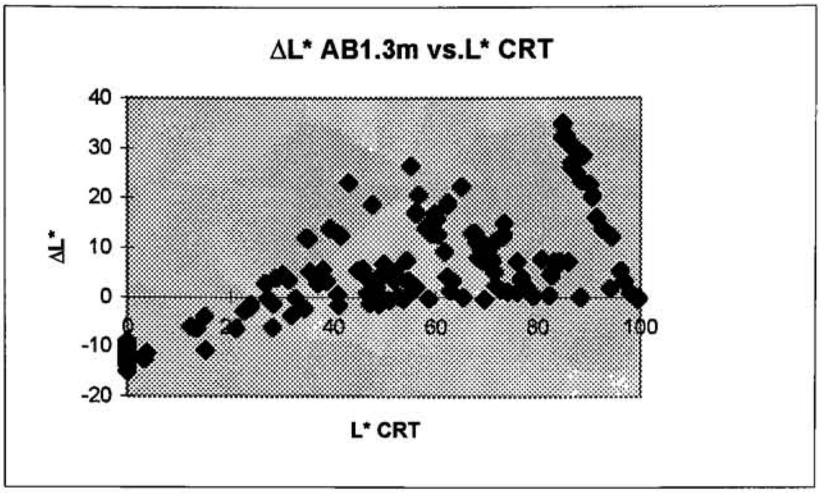


Figure 72.

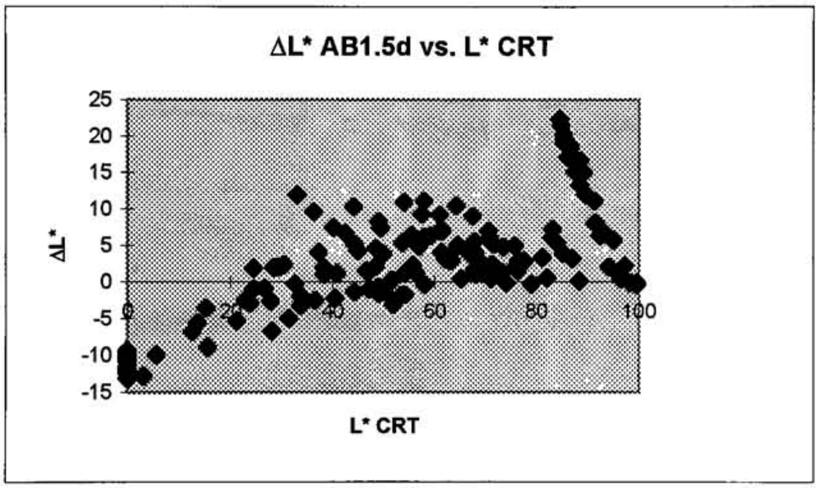


Figure 73.

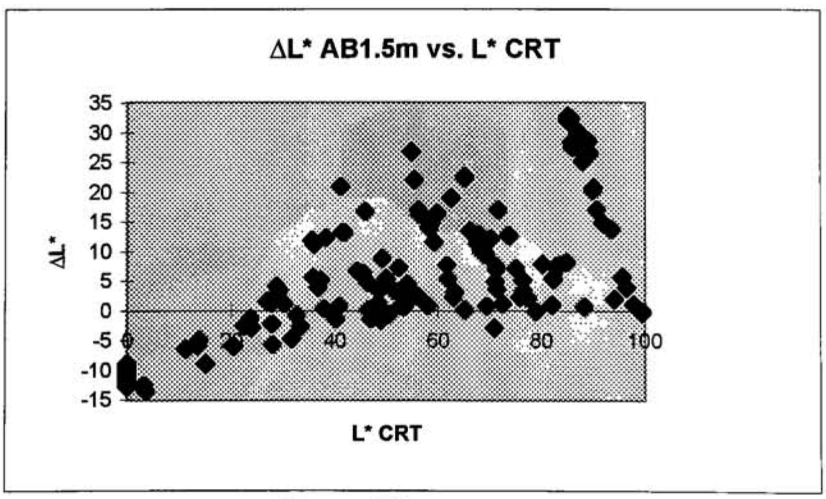


Figure 74.

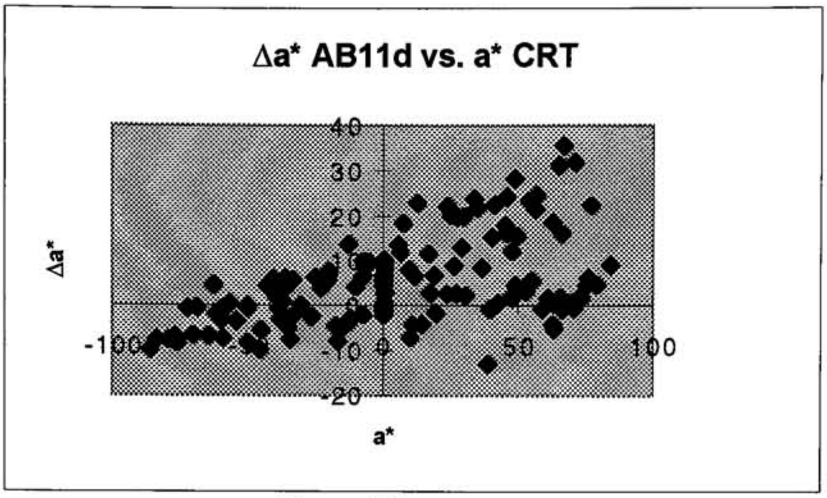


Figure 75.

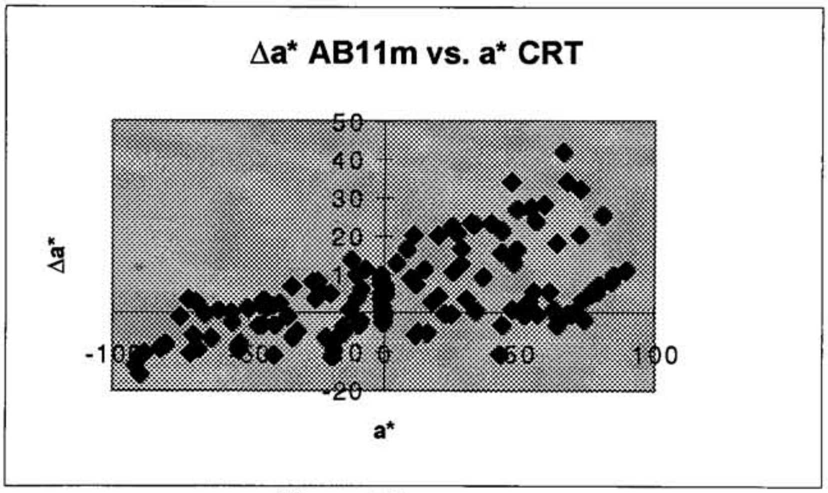


Figure 76.

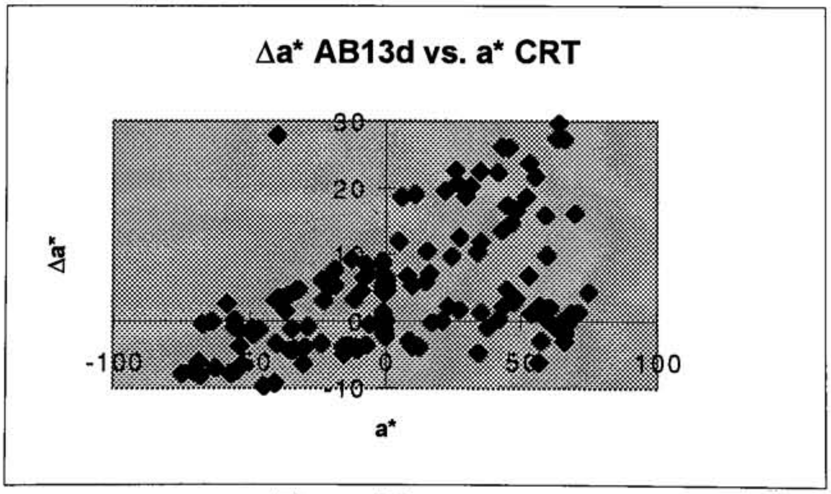


Figure 77.

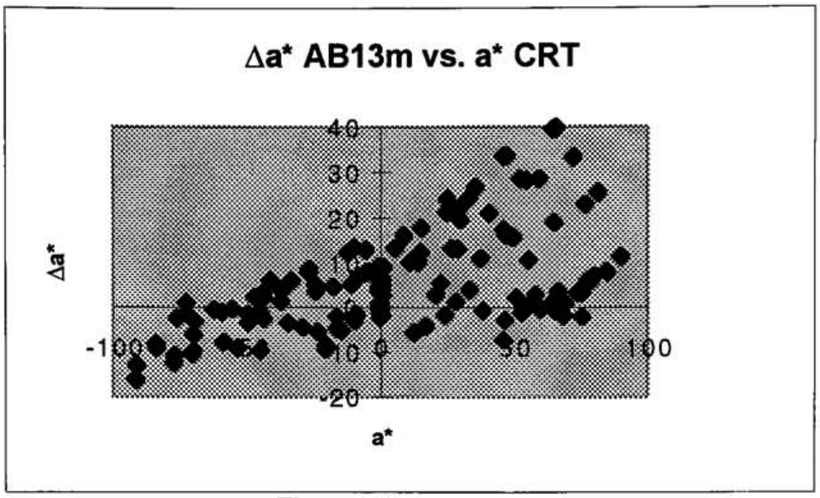


Figure 78.

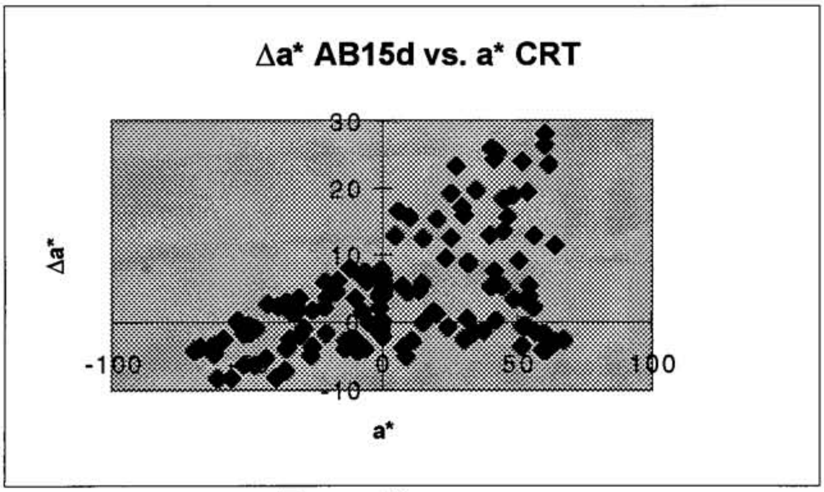


Figure 79.

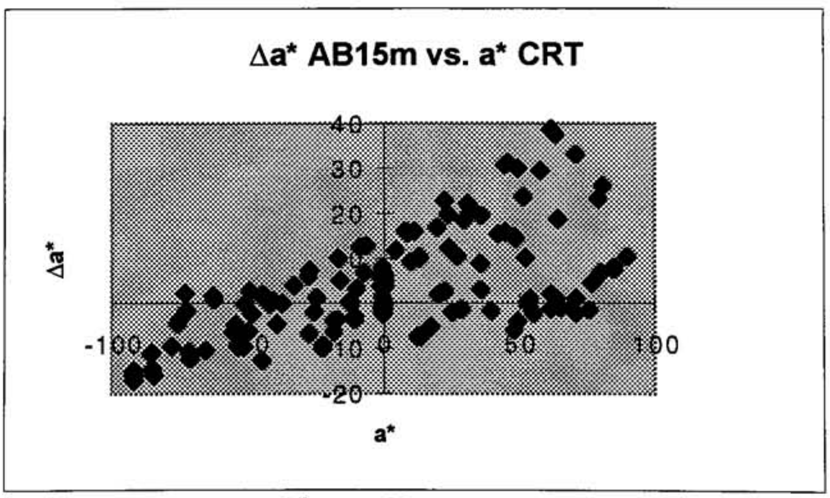


Figure 80.

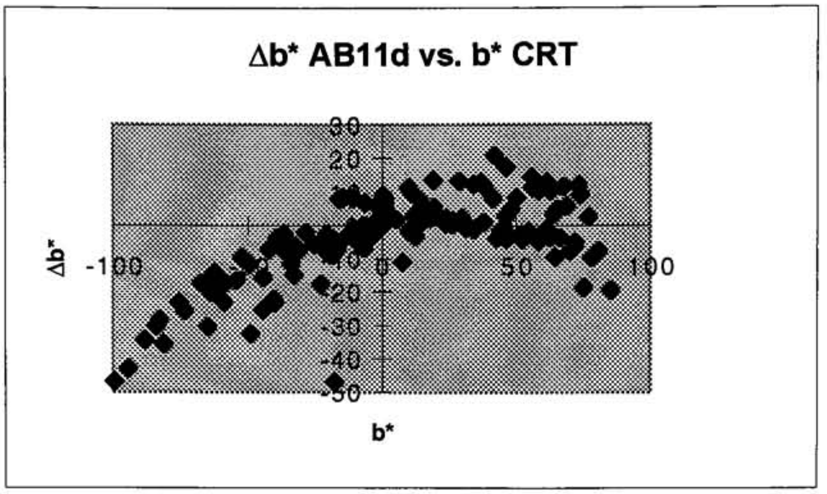


Figure 81.

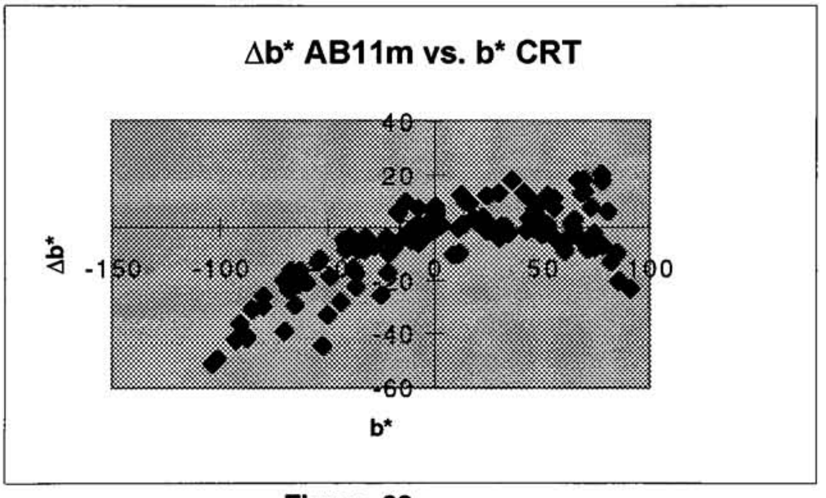


Figure 82.

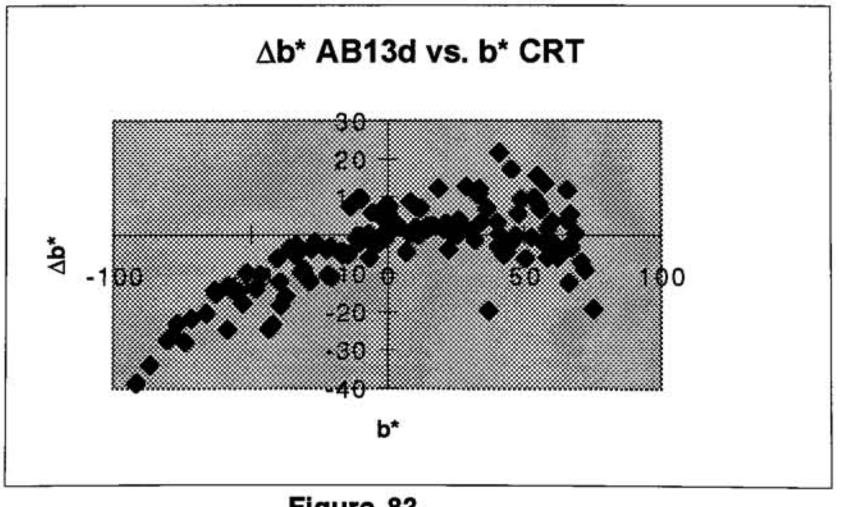


Figure 83.

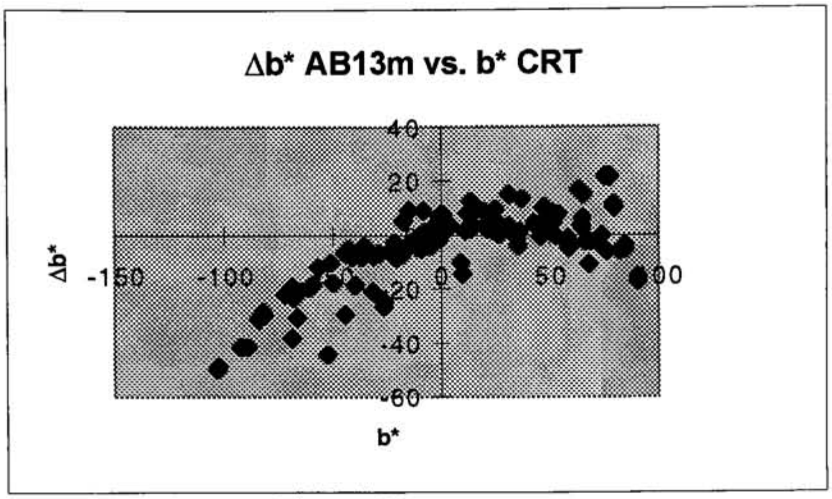


Figure 84.

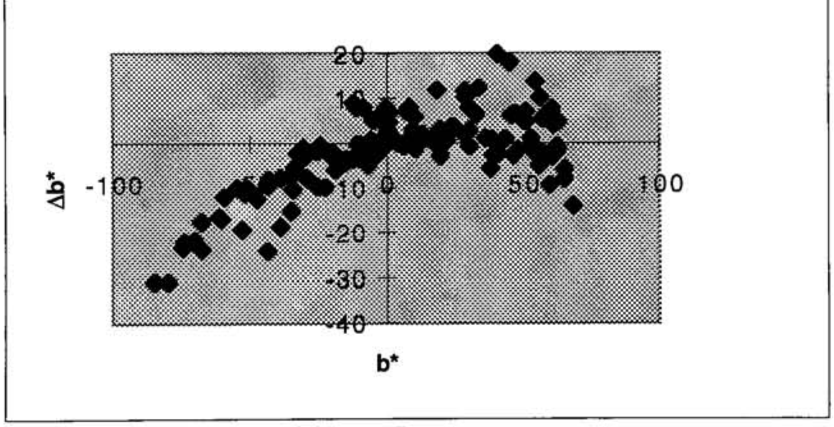


Figure 85.

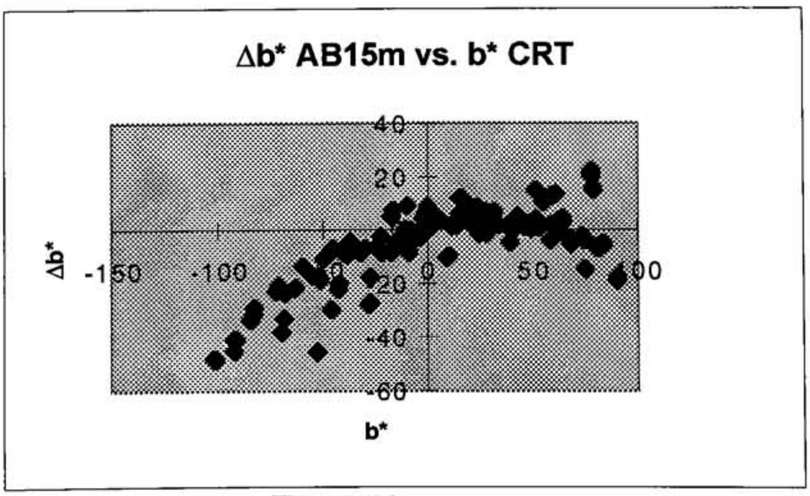


Figure 86.

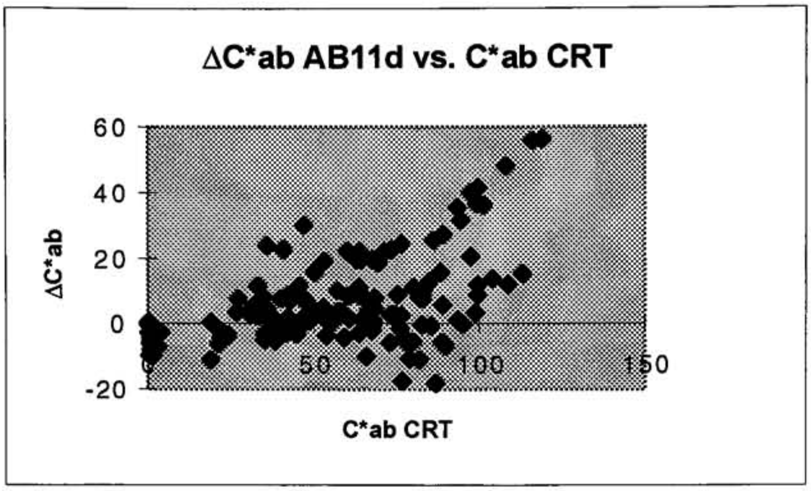


Figure 87.

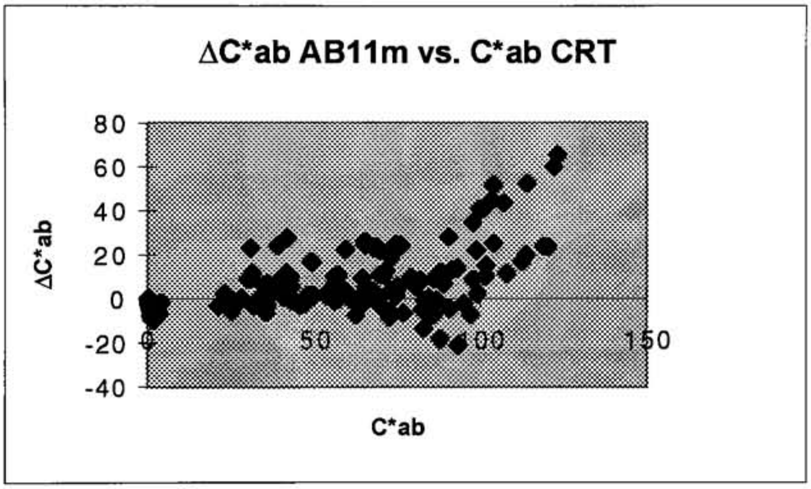


Figure 88.

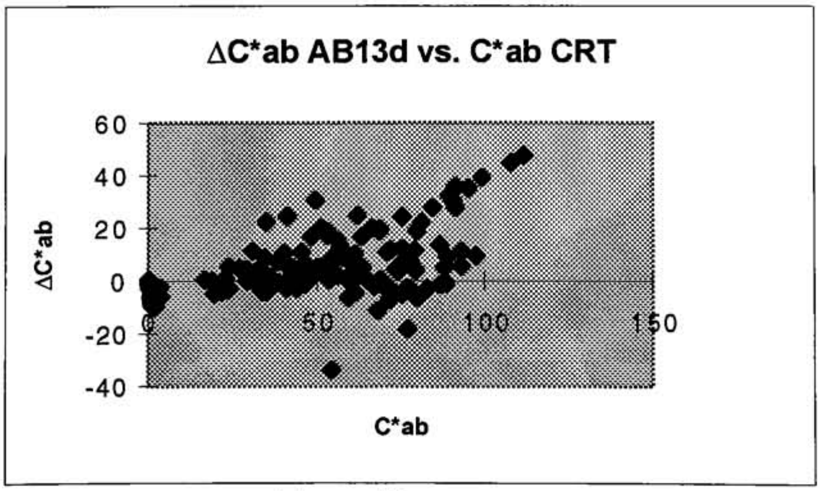


Figure 89.

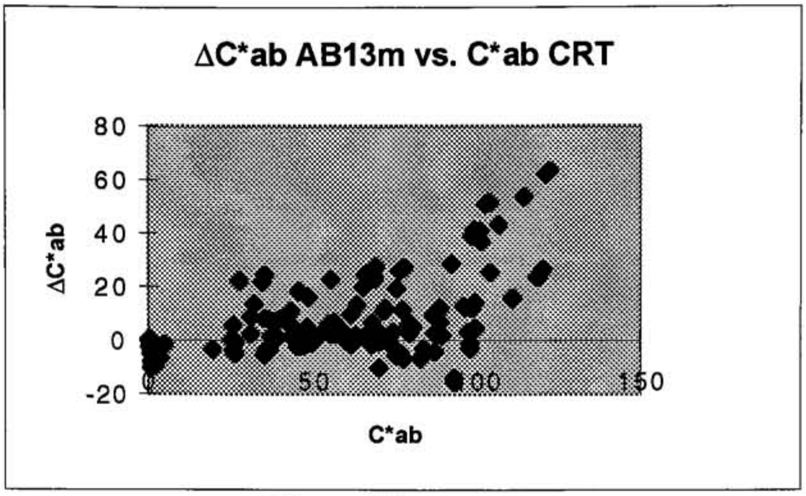


Figure 90.

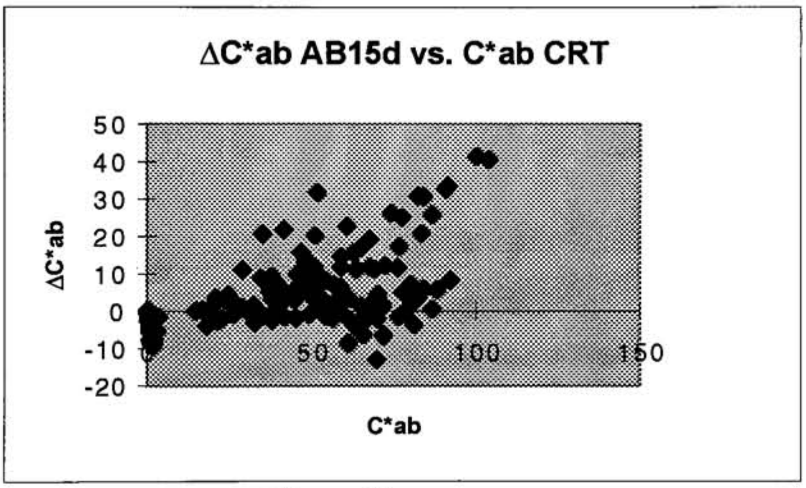


Figure 91.

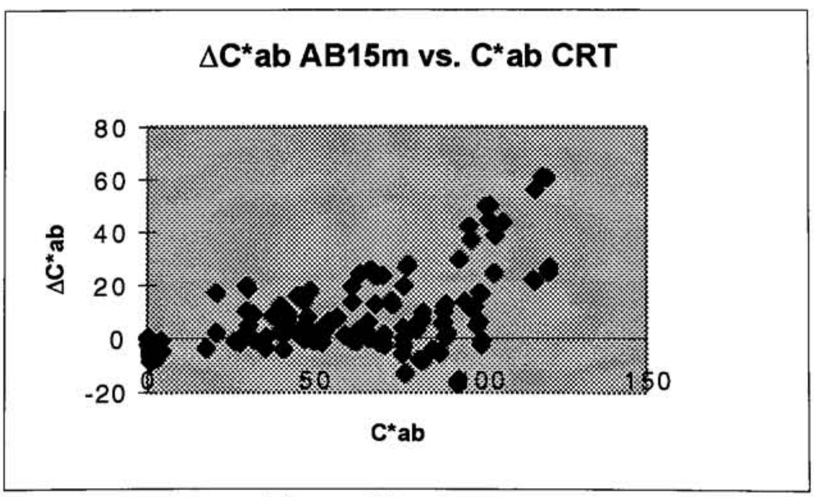


Figure 92.

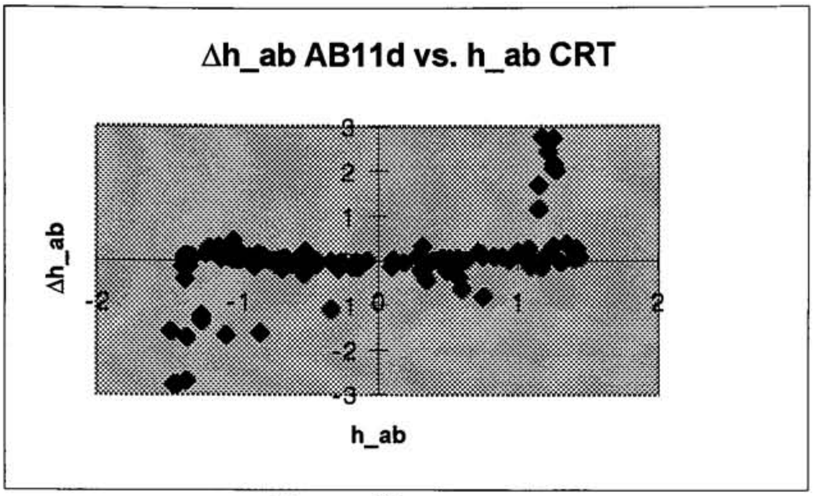


Figure 93.

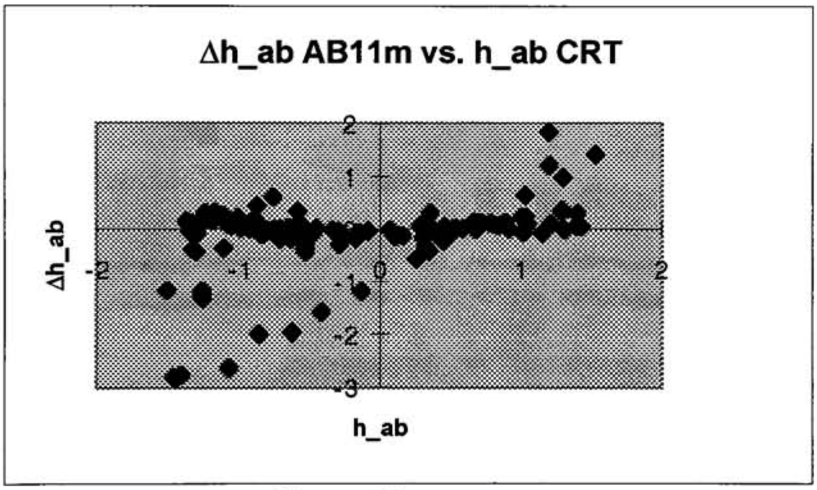


Figure 94.

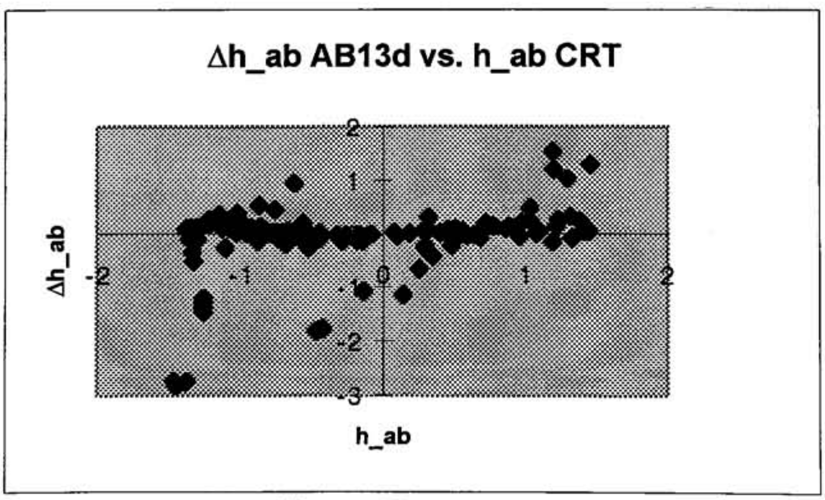


Figure 95.

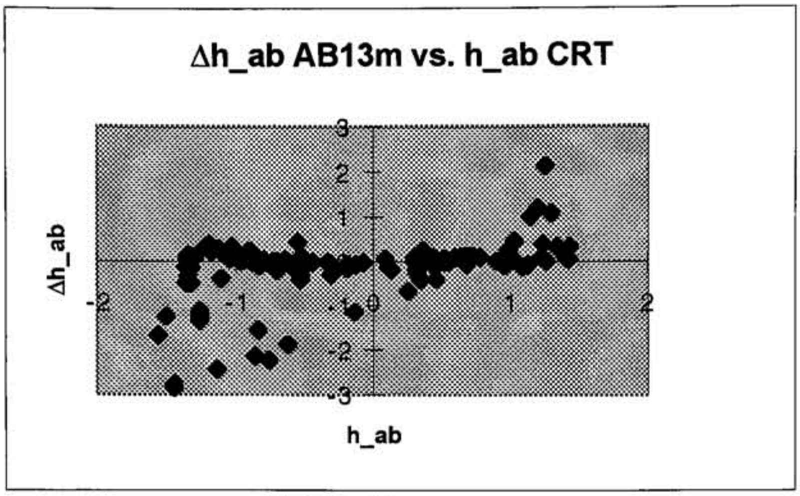


Figure 96.

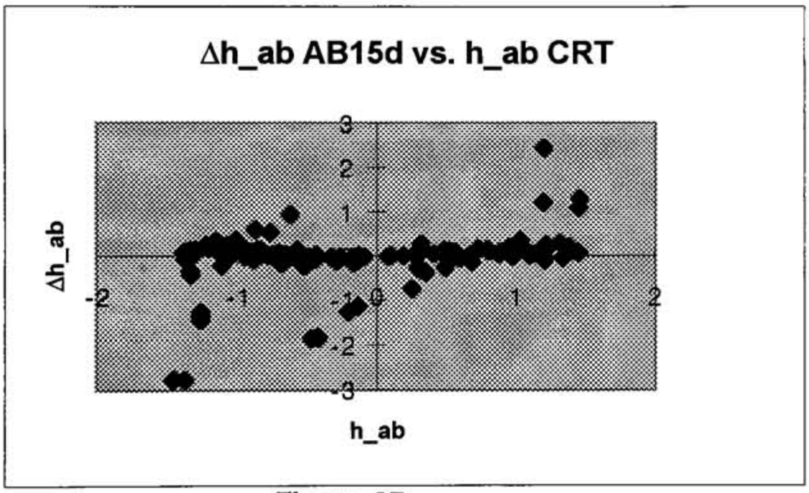


Figure 97.

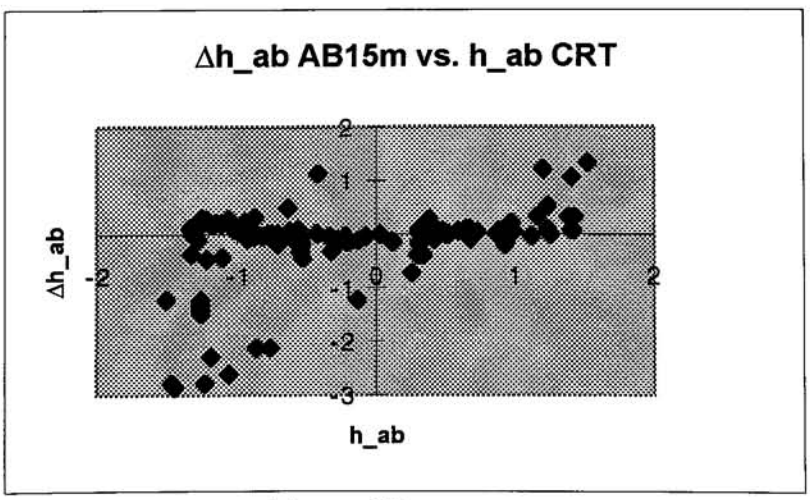


Figure 98.

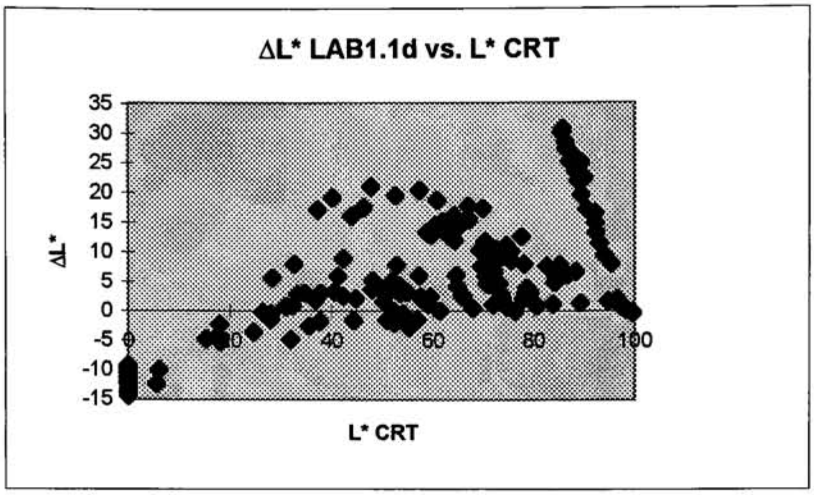


Figure 99.

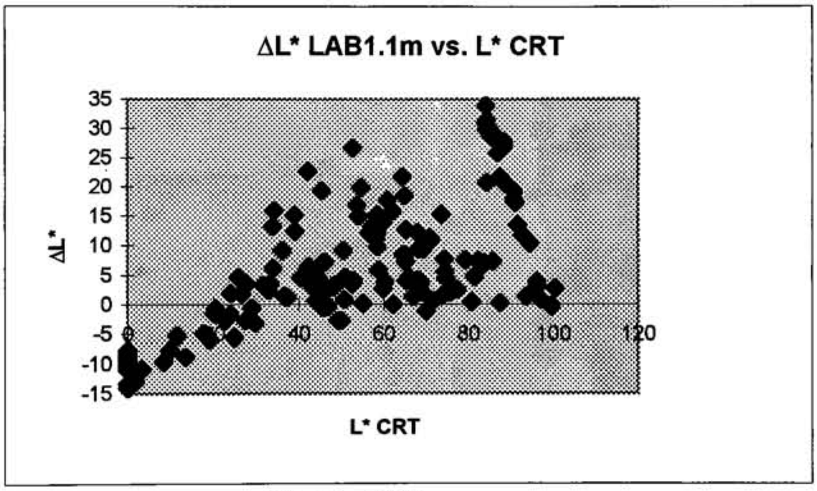


Figure 100.

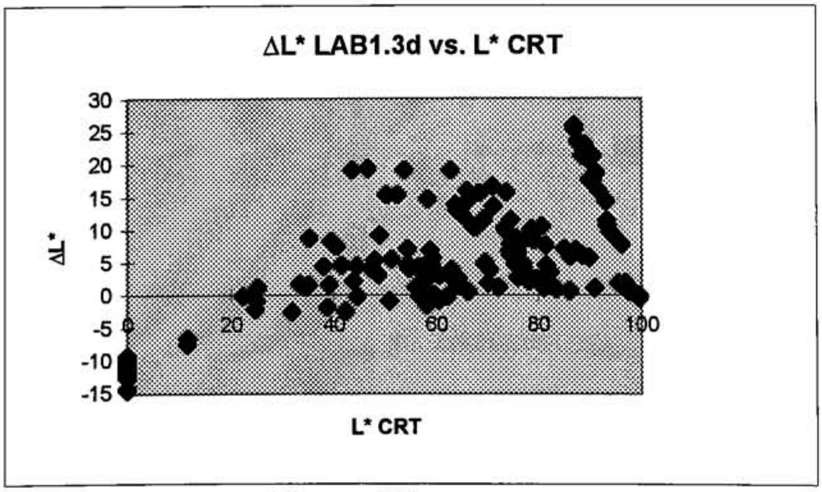


Figure 101.

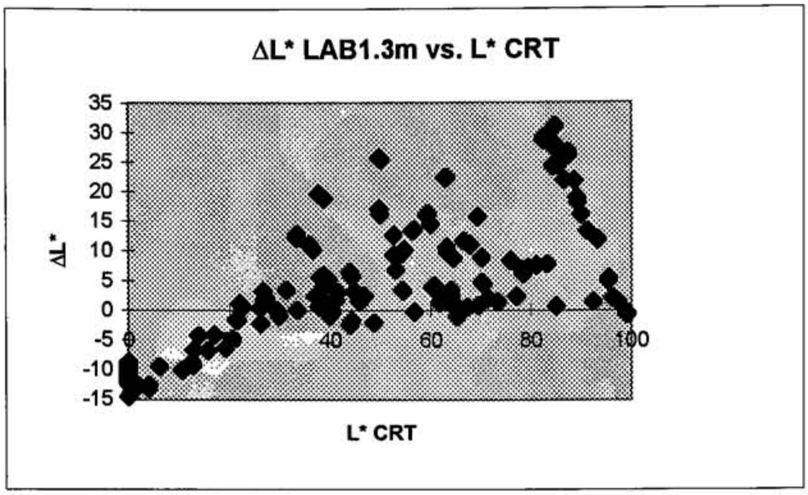


Figure 102.

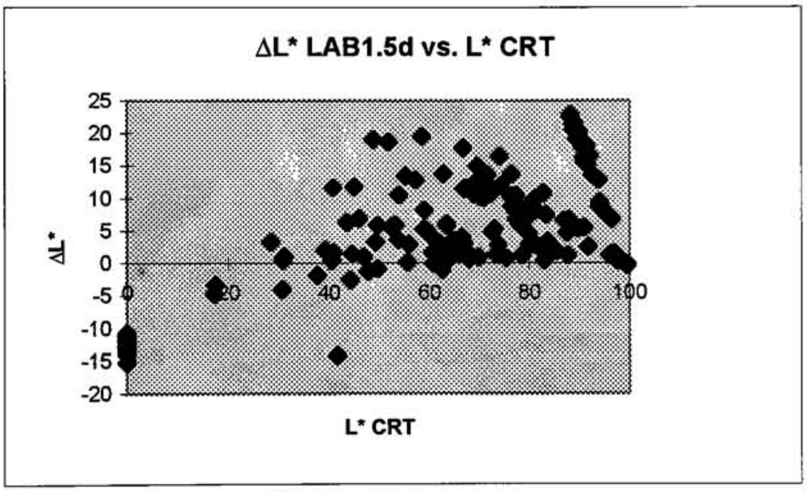


Figure 103.

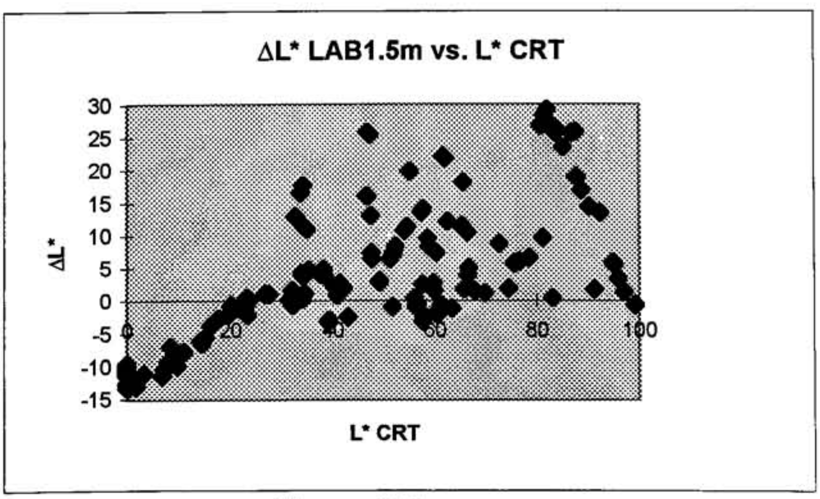


Figure 104.

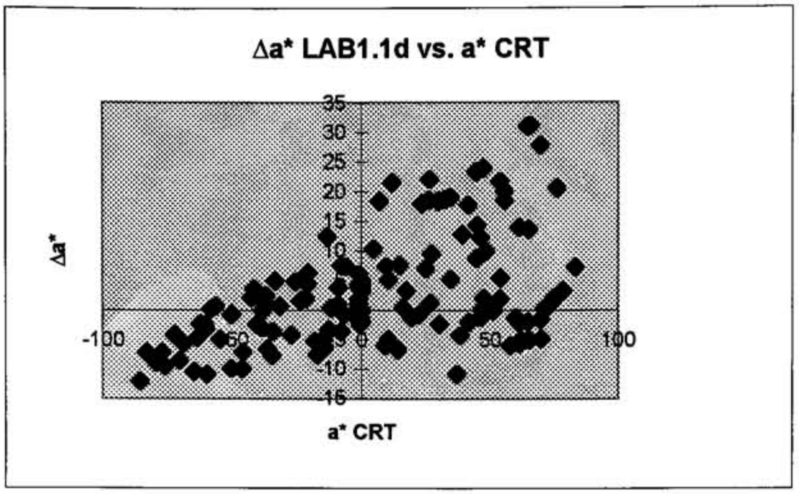


Figure 105.

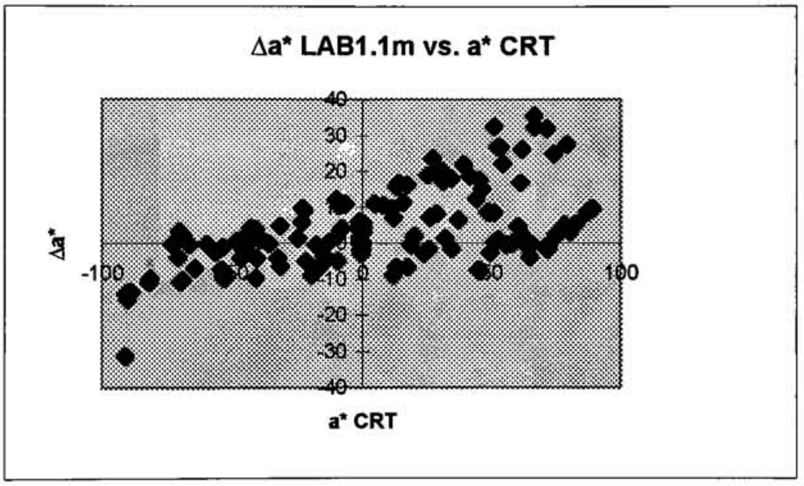


Figure 106.

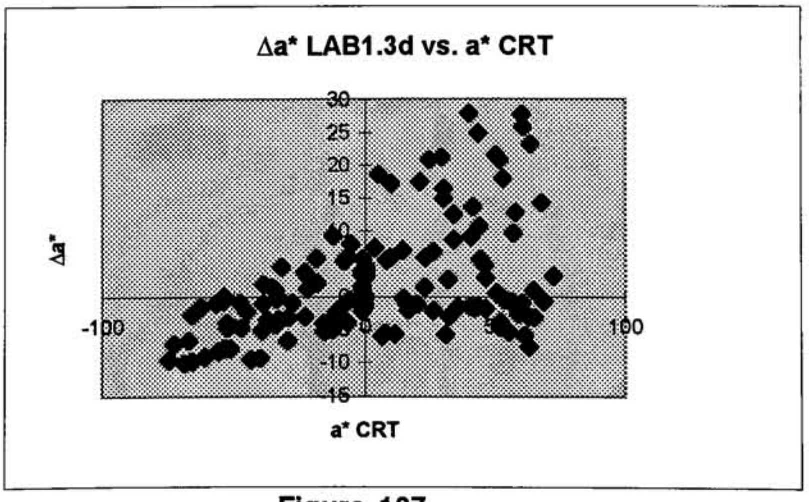


Figure 107.

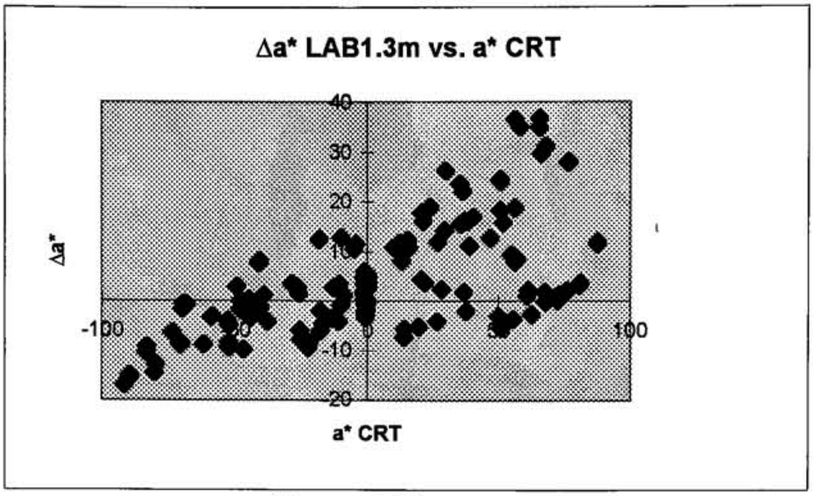


Figure 108.

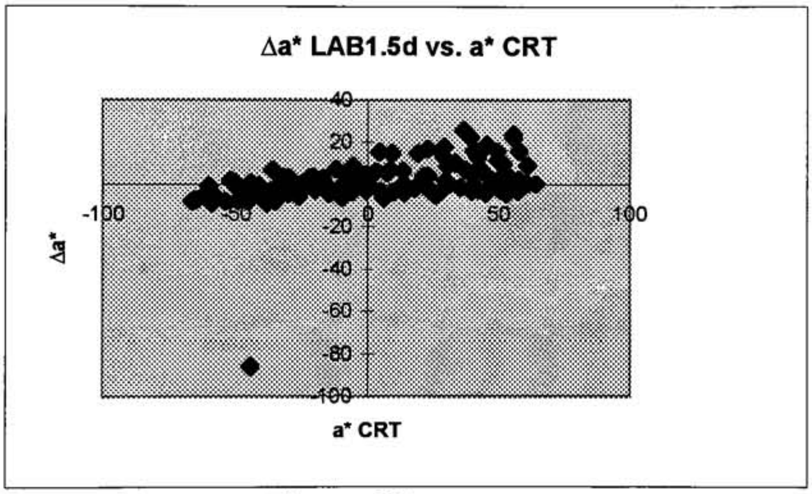


Figure 109.

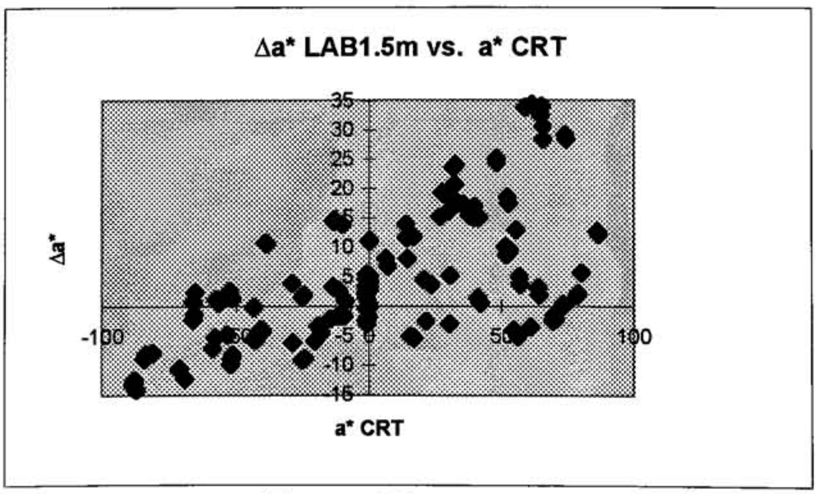


Figure 110.

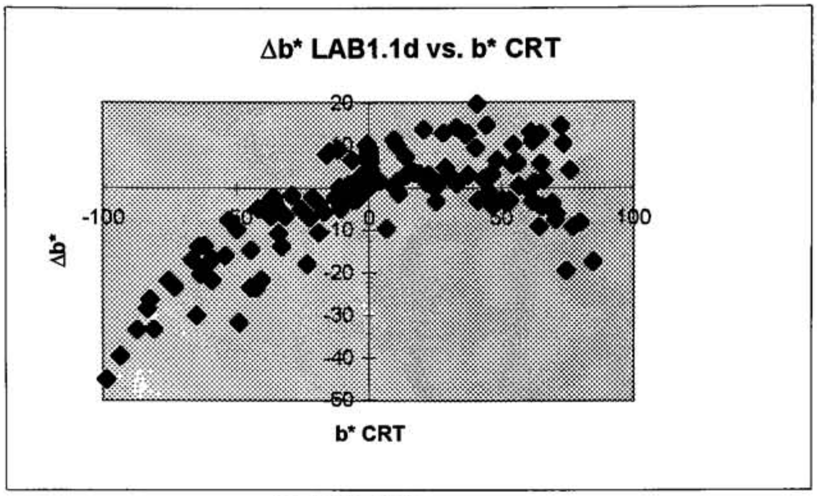


Figure 111.

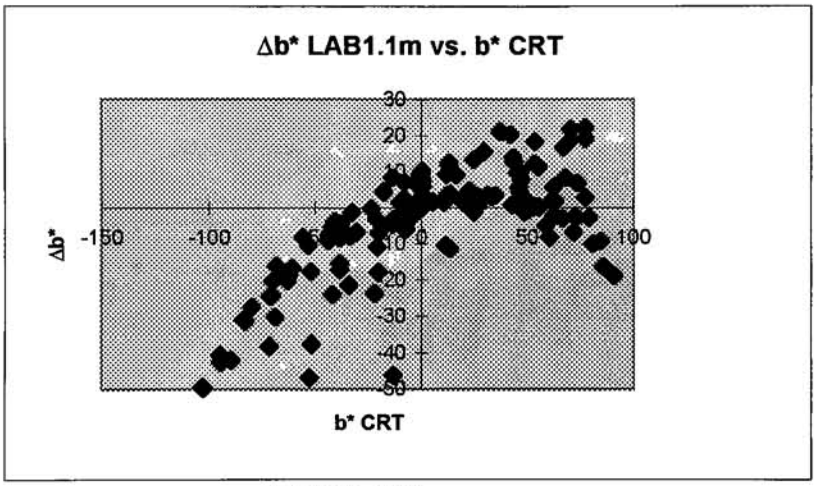


Figure 112.

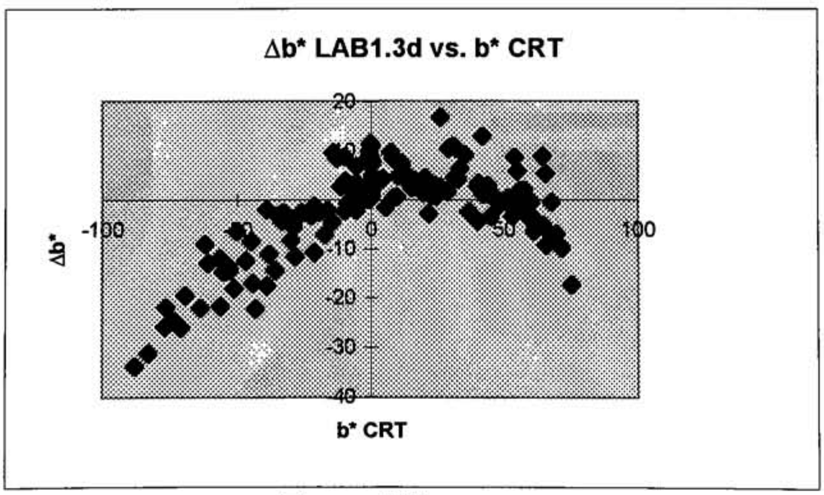


Figure 113.

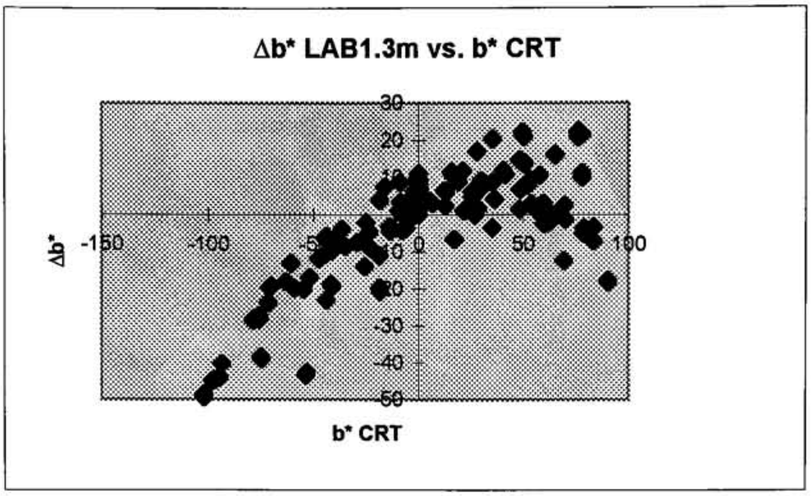


Figure 114.

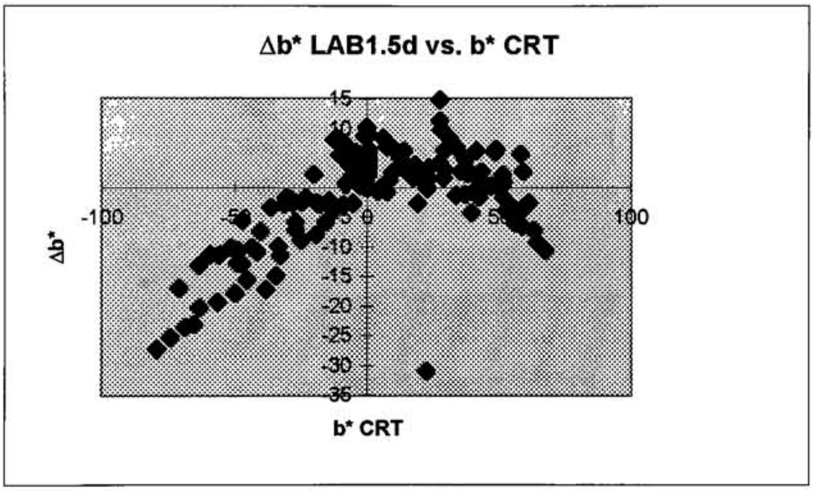


Figure 115.

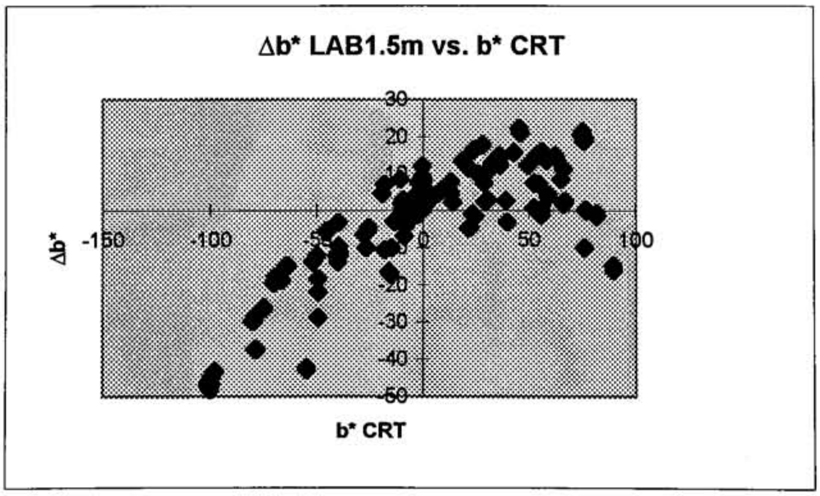


Figure 116.

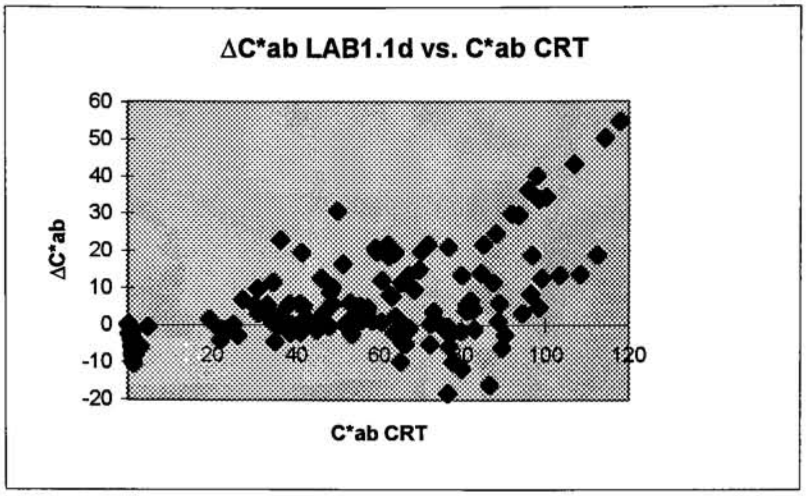


Figure 117.

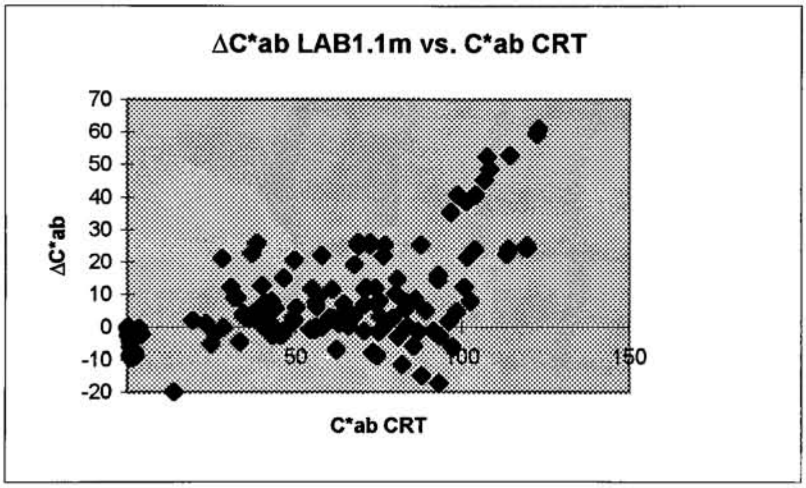


Figure 118.

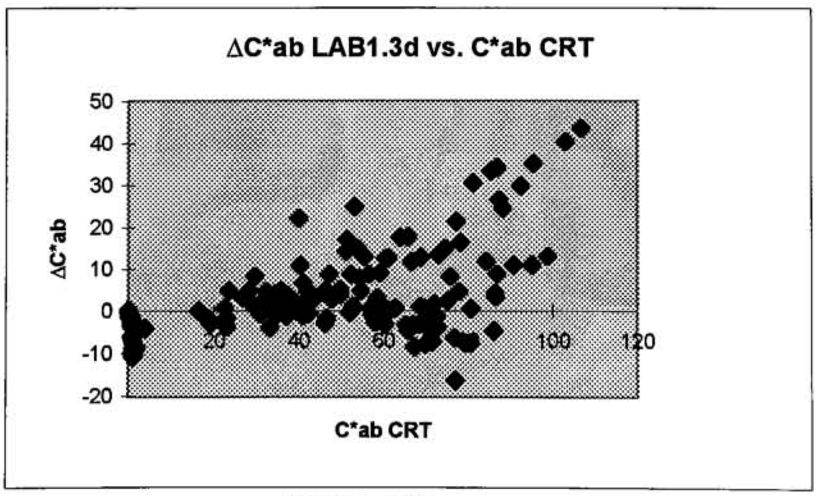


Figure 119.

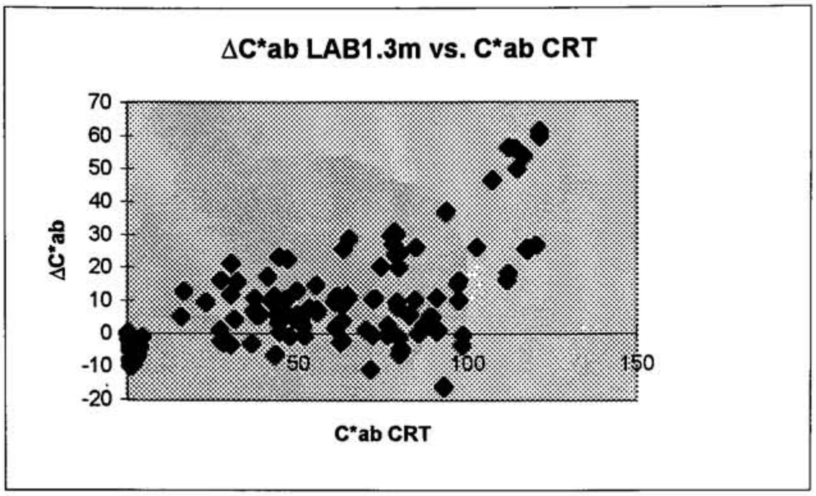


Figure 120.

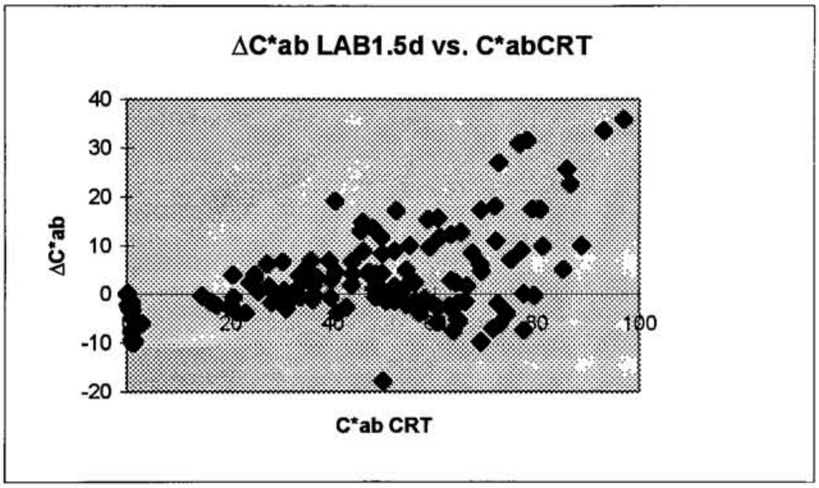


Figure 121.

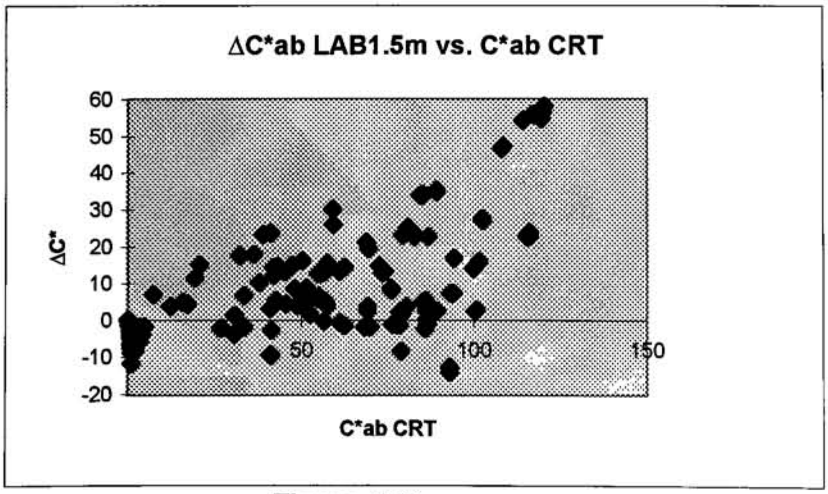


Figure 122.

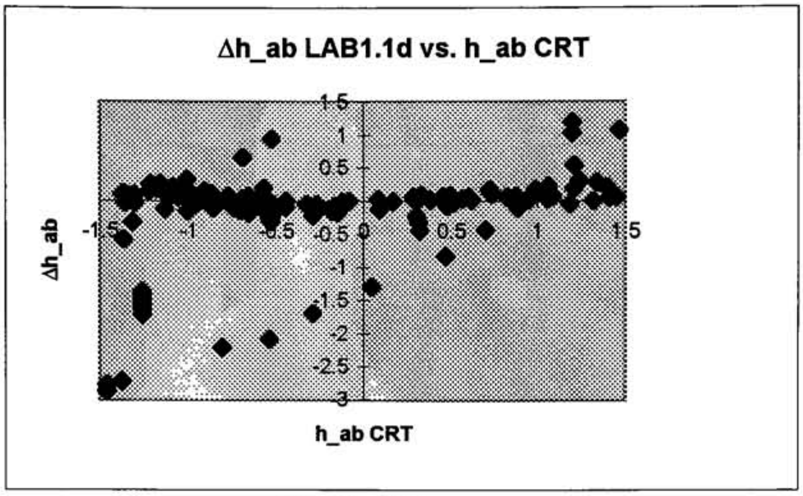


Figure 123.

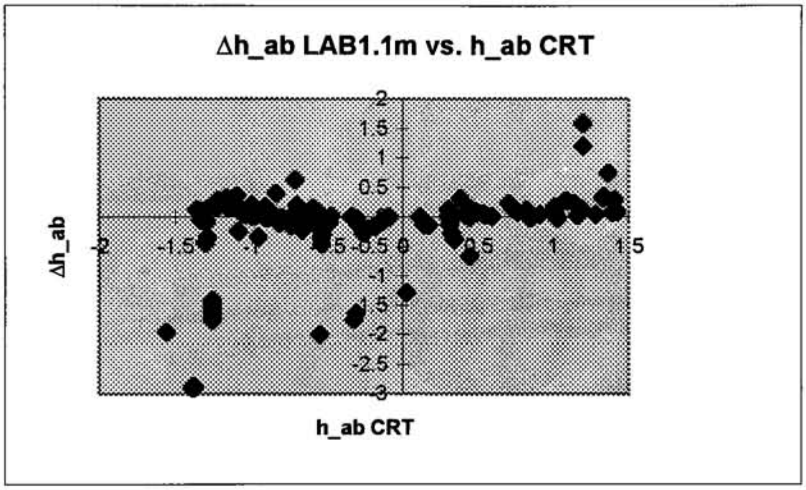


Figure 124.

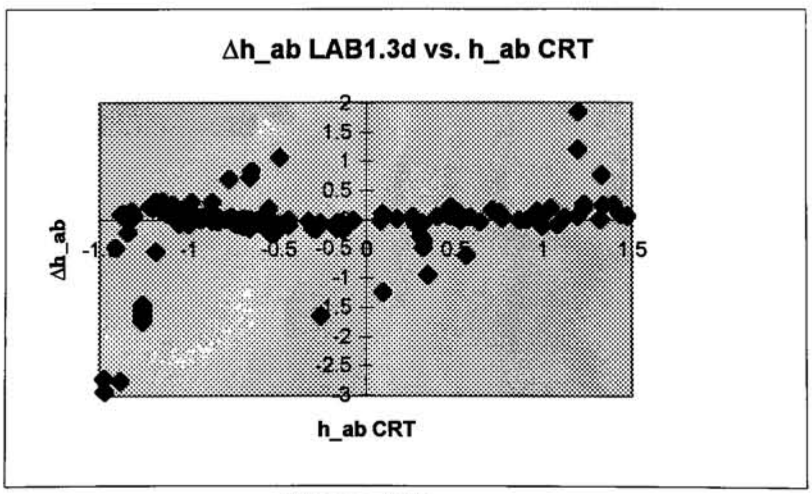


Figure 125.

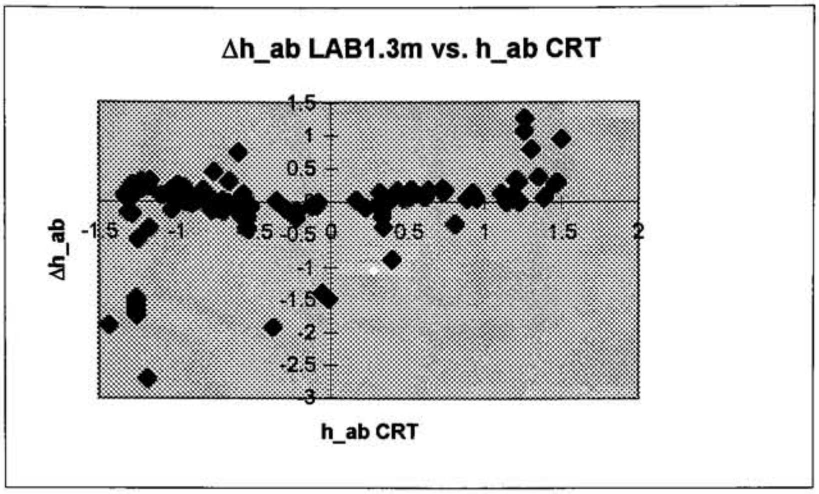


Figure 126.

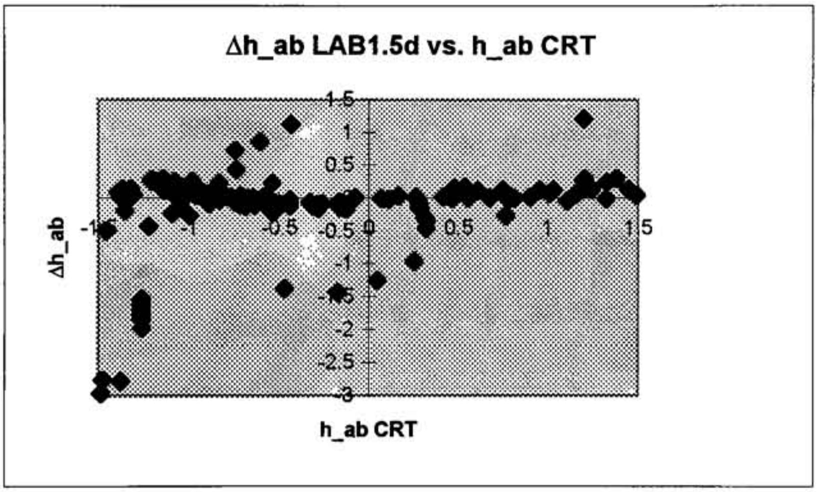


Figure 127.

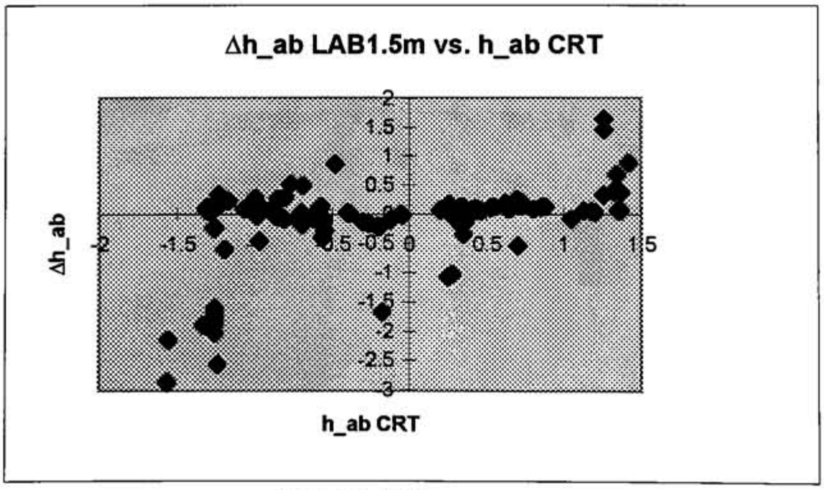


Figure 128.

Appendix 3.

The histograms of the ΔE_{94} and MCDM for the 5x5x5 target are included here.

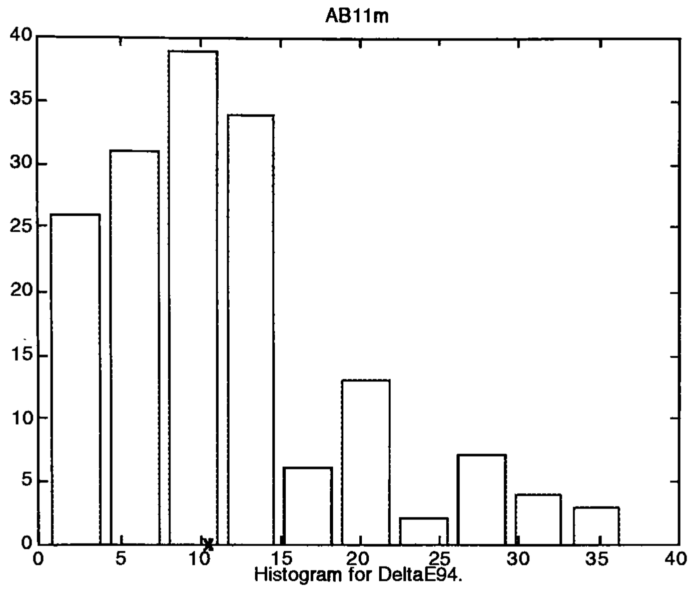


Figure 1.

med =10.48
maxim = 36.64

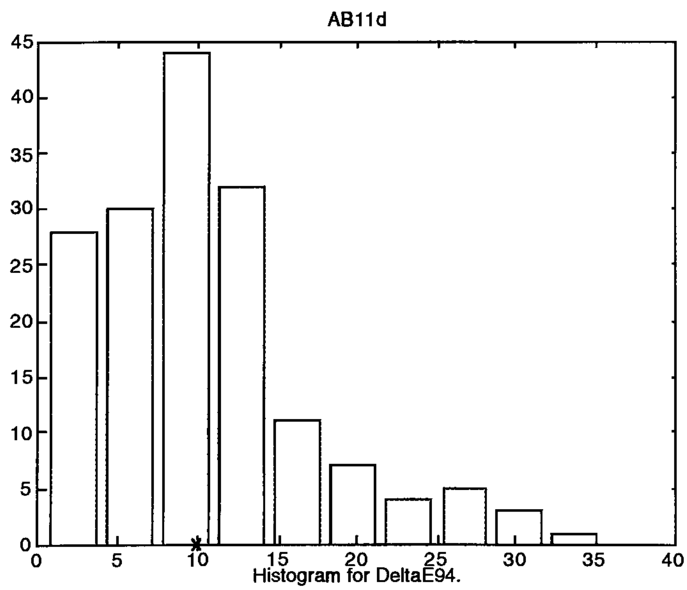


Figure 2.
med = 9.87
maxim = 35.47

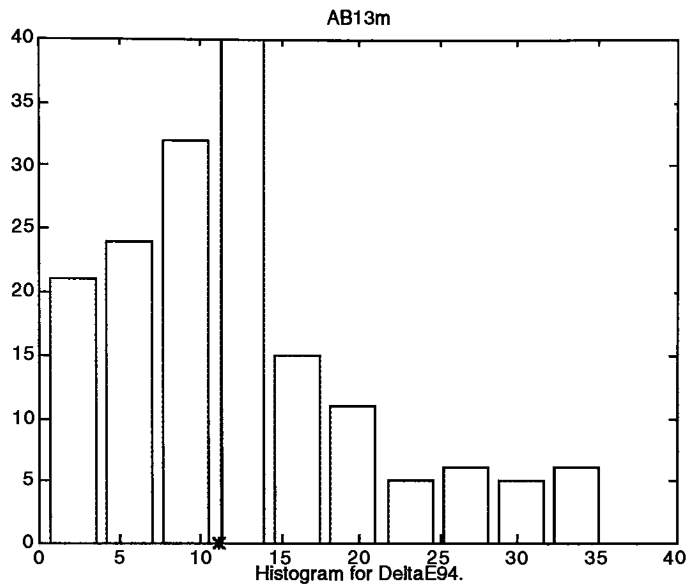


Figure 3.
med =11.11
maxim =35.42

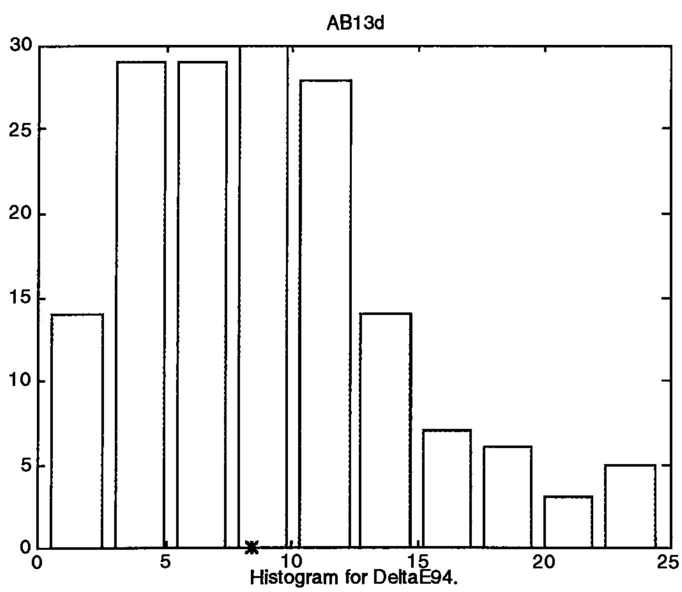


Figure 4.
med = 8.41
maxim = 24.60

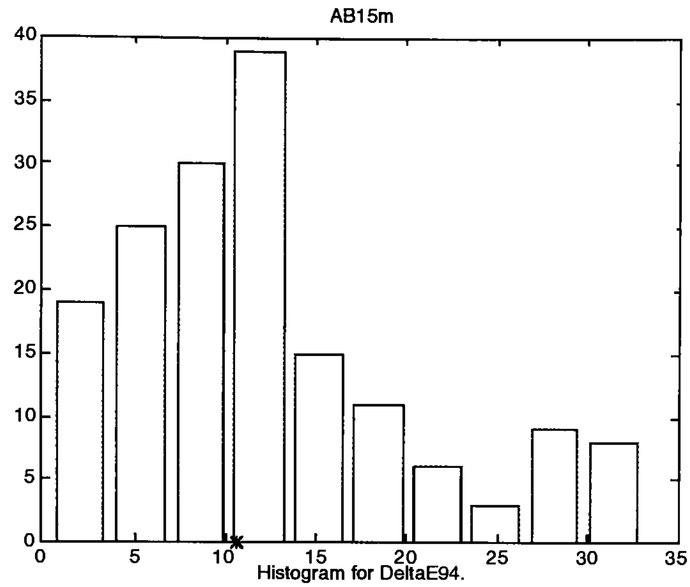


Figure 5.
med =10.57
maxim =33.05

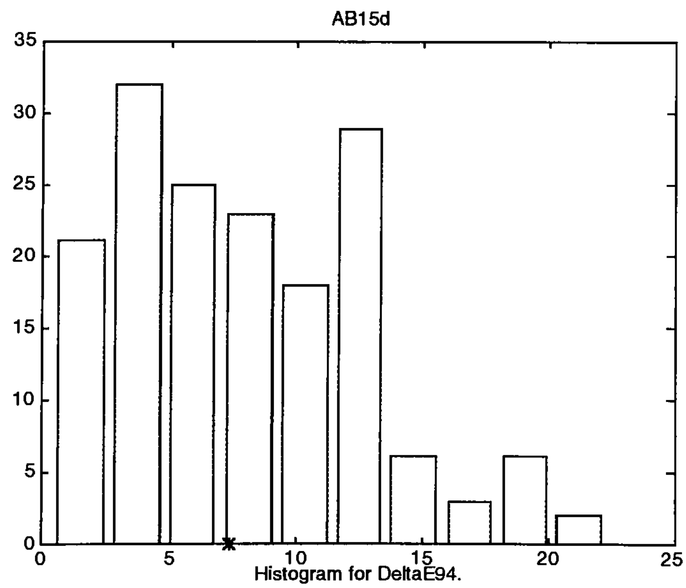


Figure6.
med =7.34
maxim = 22.32

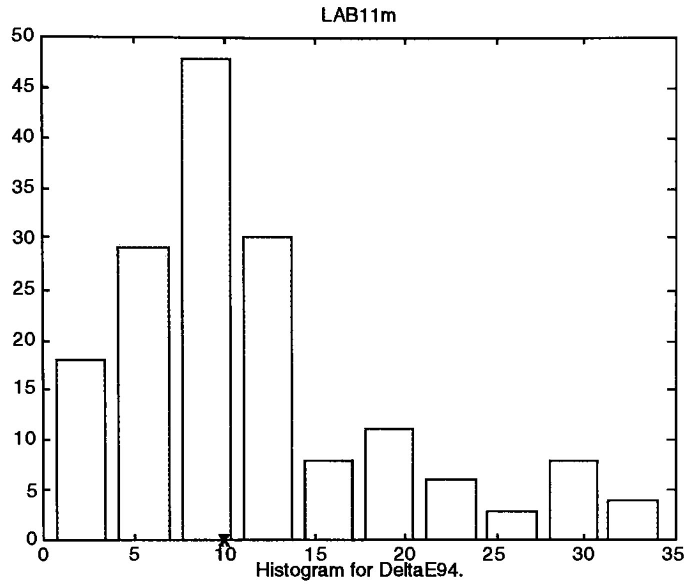


Figure 7.
med =9.93
maxim =34.33

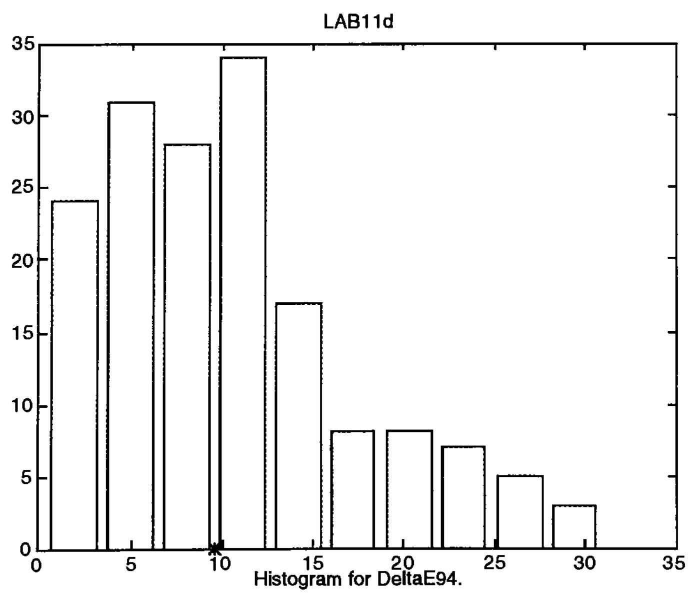


Figure 8.
med = 9.55
maxim = 30.90

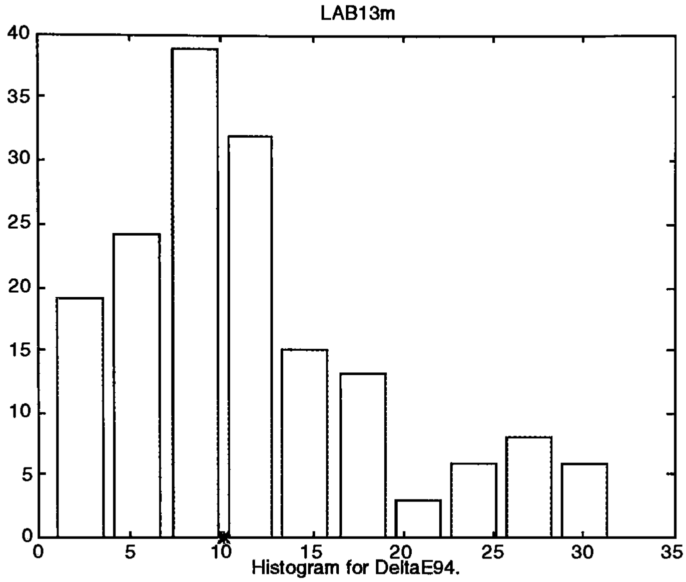


Figure 9.
med =10.11
maxim =31.57

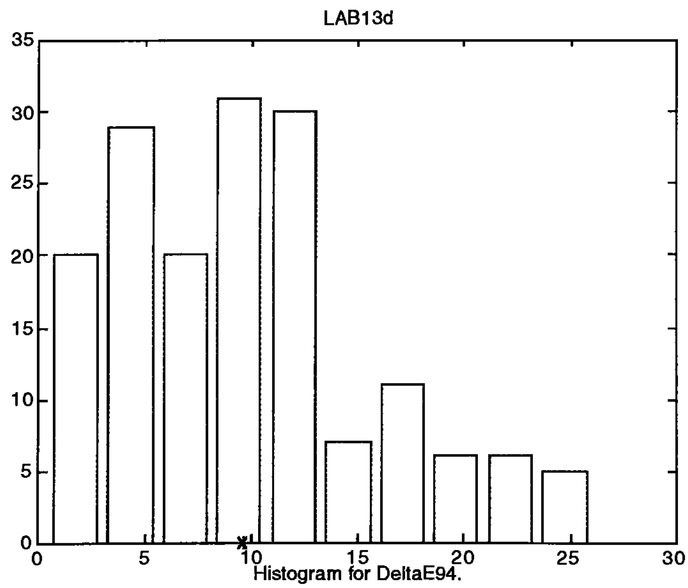


Figure 10.
med =9.51
maxim =25.99

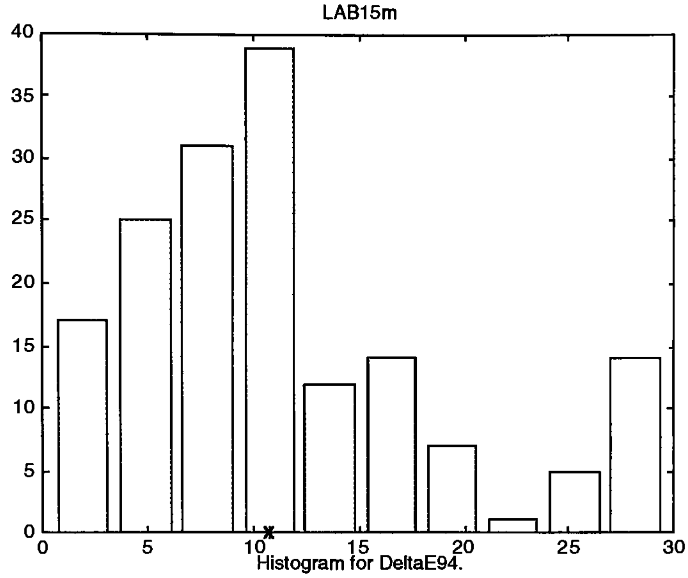


Figure 11.
med =10.68
maxim = 29.63

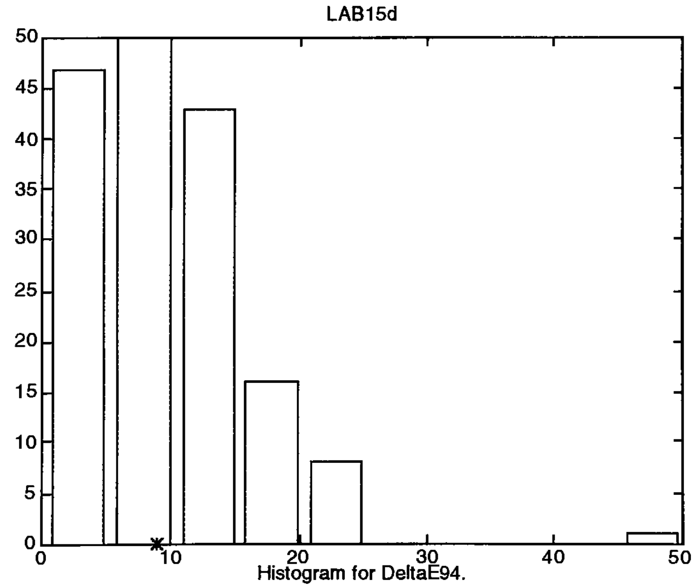


Figure 12.
med = 8.84
maxim = 50.24

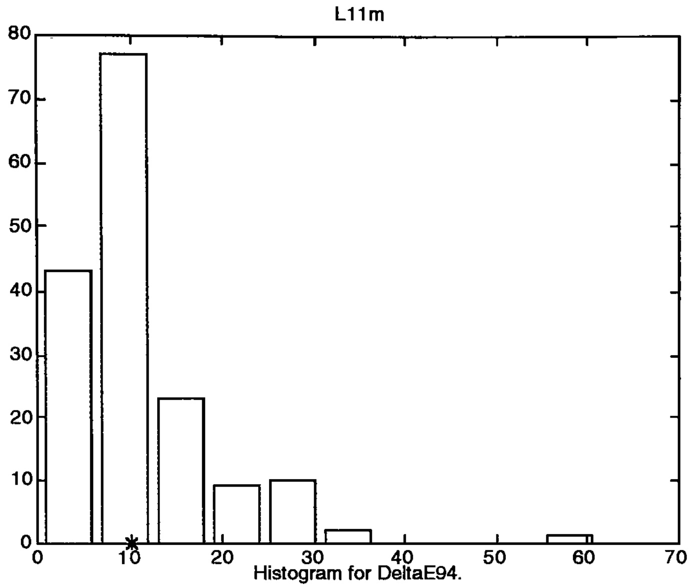


Figure 13.
med = 10.10
maxim = 61.22

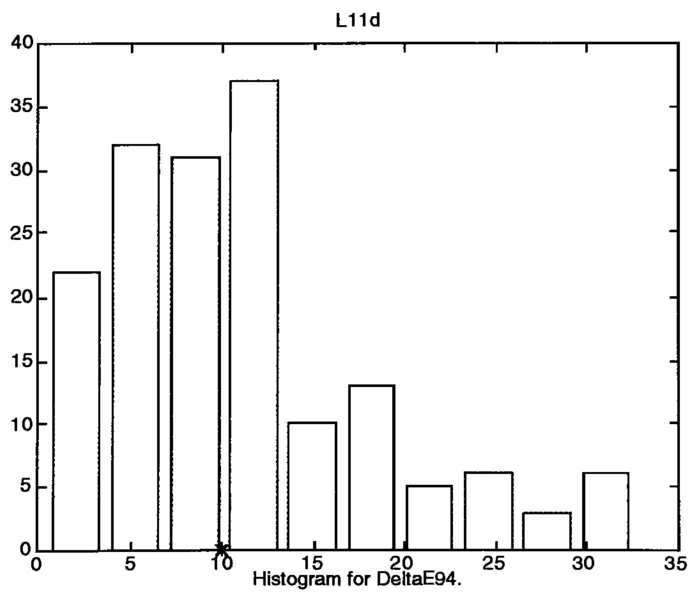


Figure 14.
med =10.02
maxim =32.64

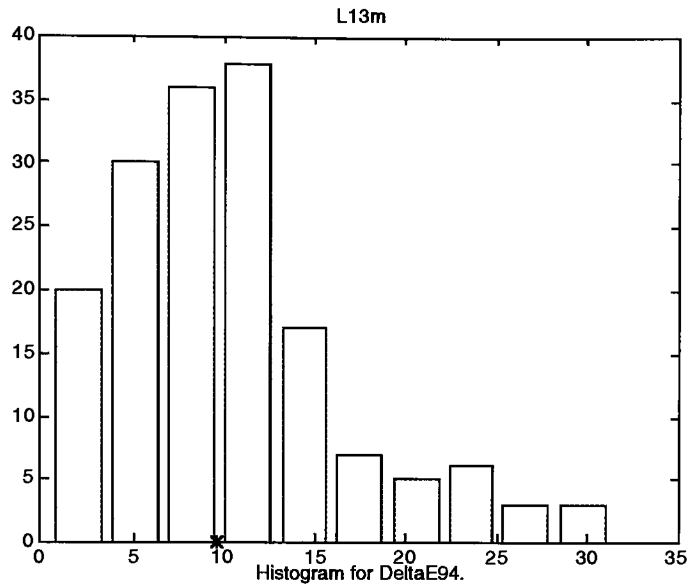


Figure 15.
med =9.57
maxim =31.23

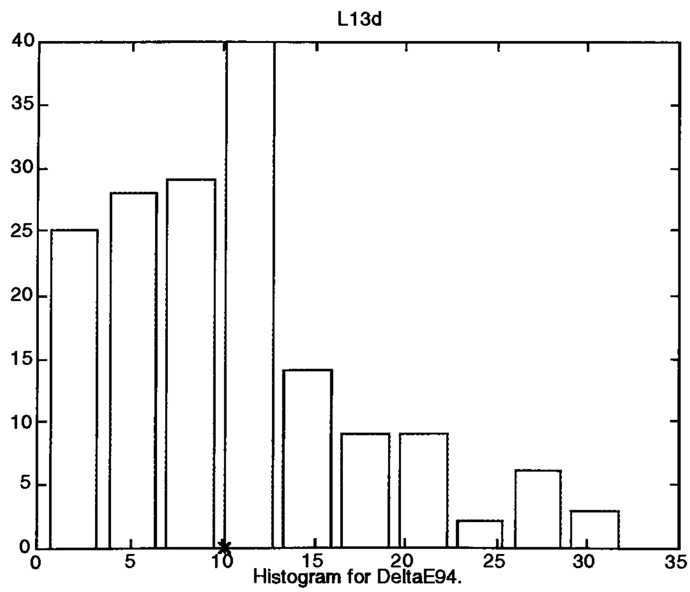


Figure 16.
med =10.05
maxim =32.06

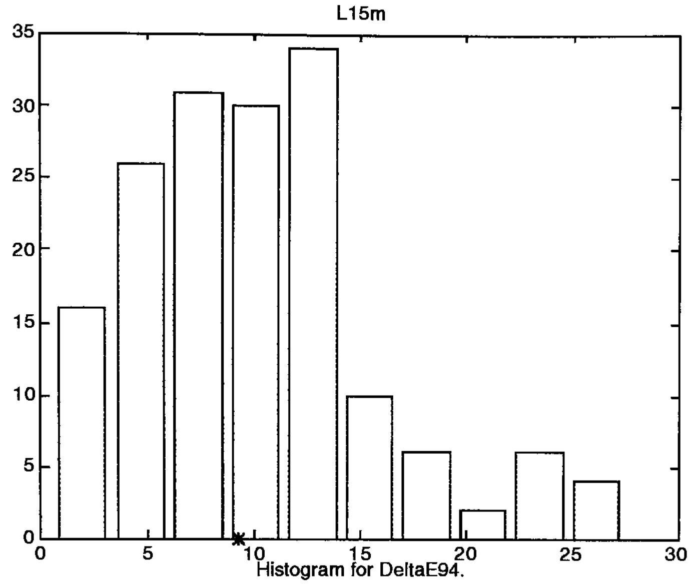


Figure 17.
med =9.24
maxim = 27.47

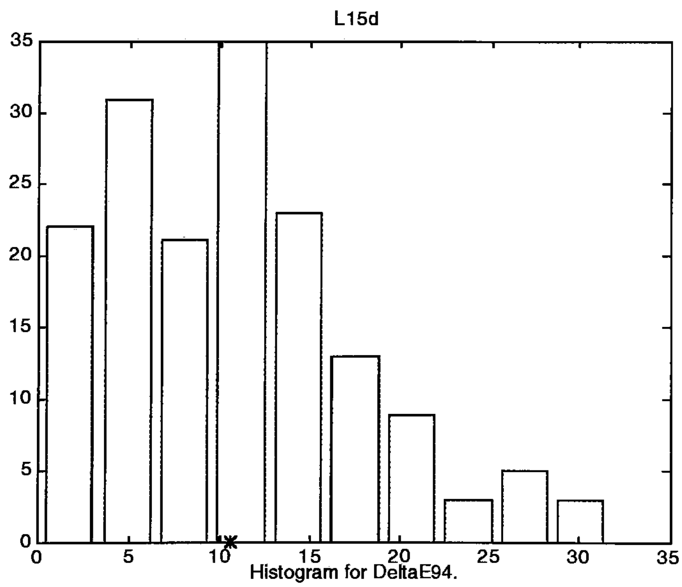


Figure 18.
med =10.58
maxim = 31.55

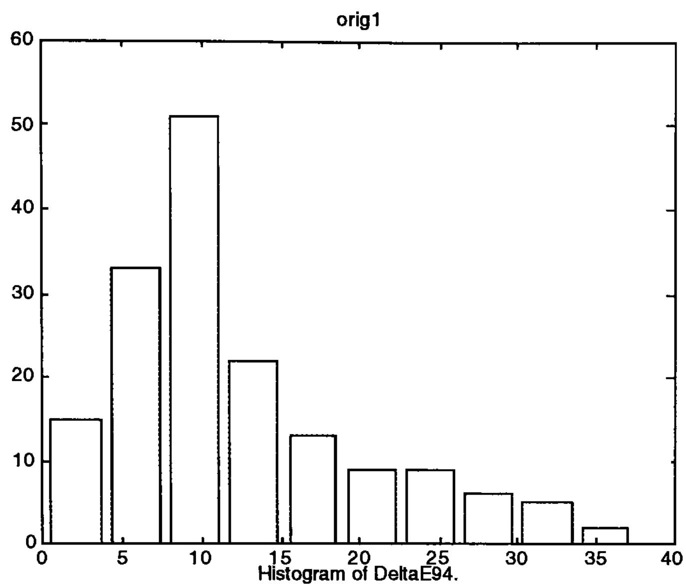


Figure 19.
med =10.63
maxim =37.48

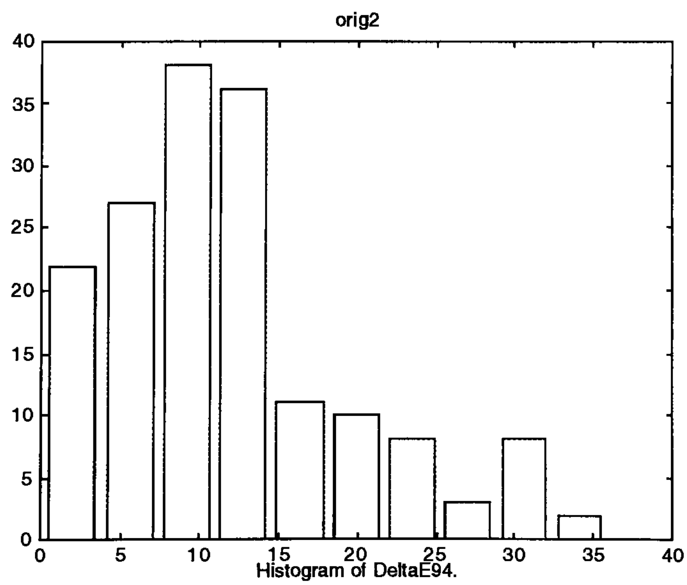


Figure 20.
med =10.19
maxim =35.92

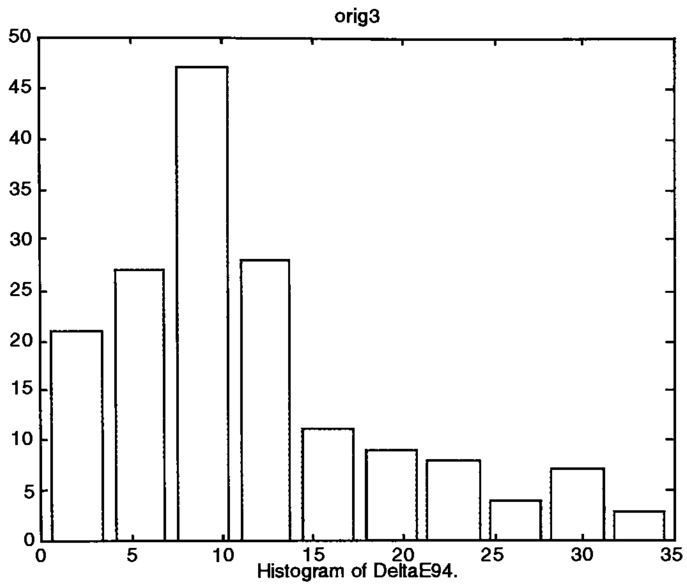


Figure 21.
med =9.52
maxim =34.79

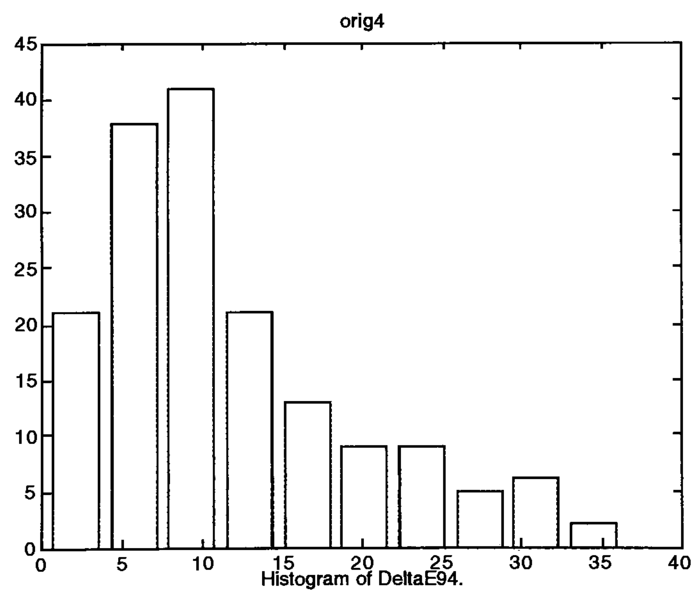


Figure 22.
med =9.00
maxim =36.33

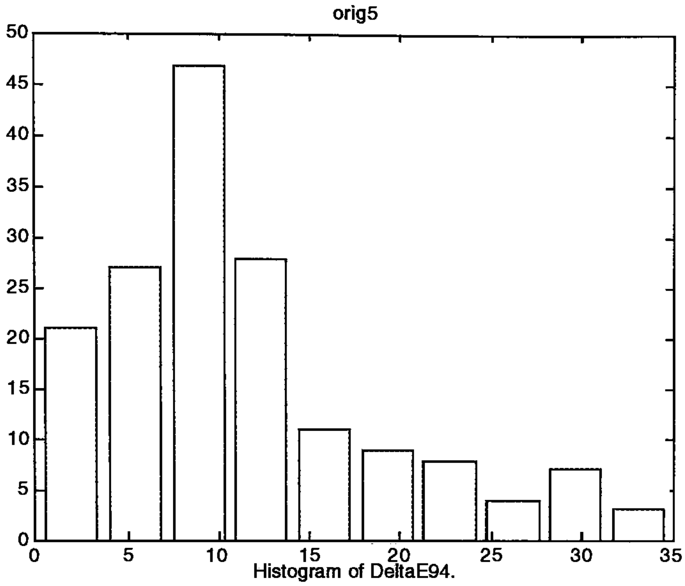


Figure 23.
med = 9.53
maxim = 34.83

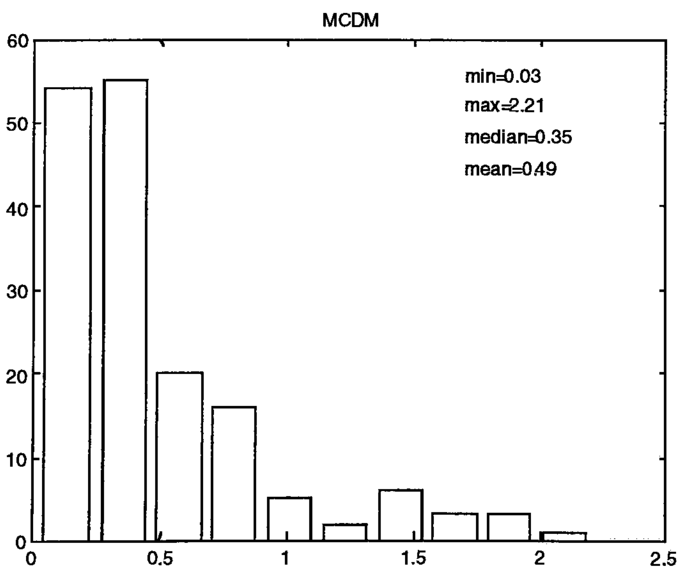
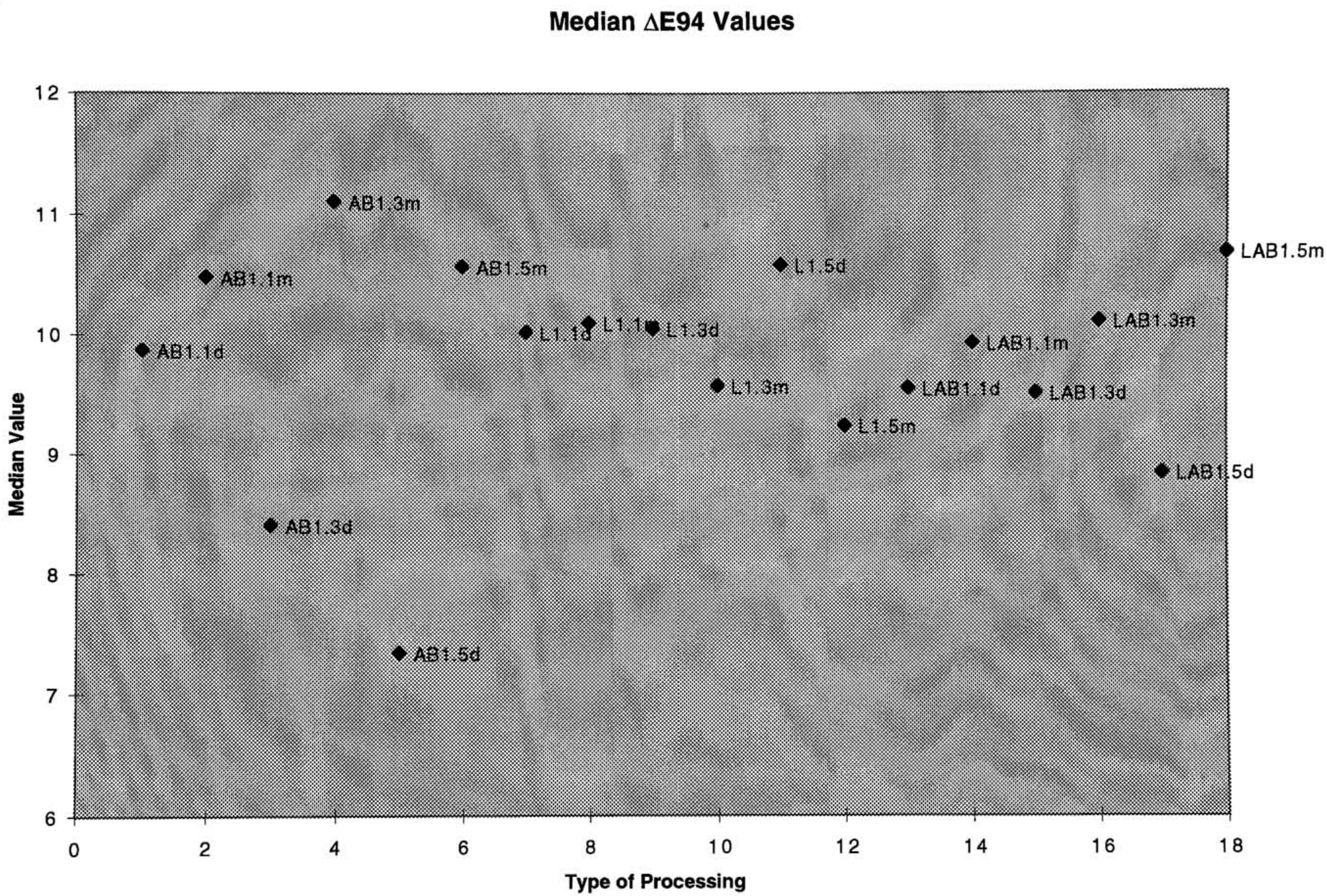


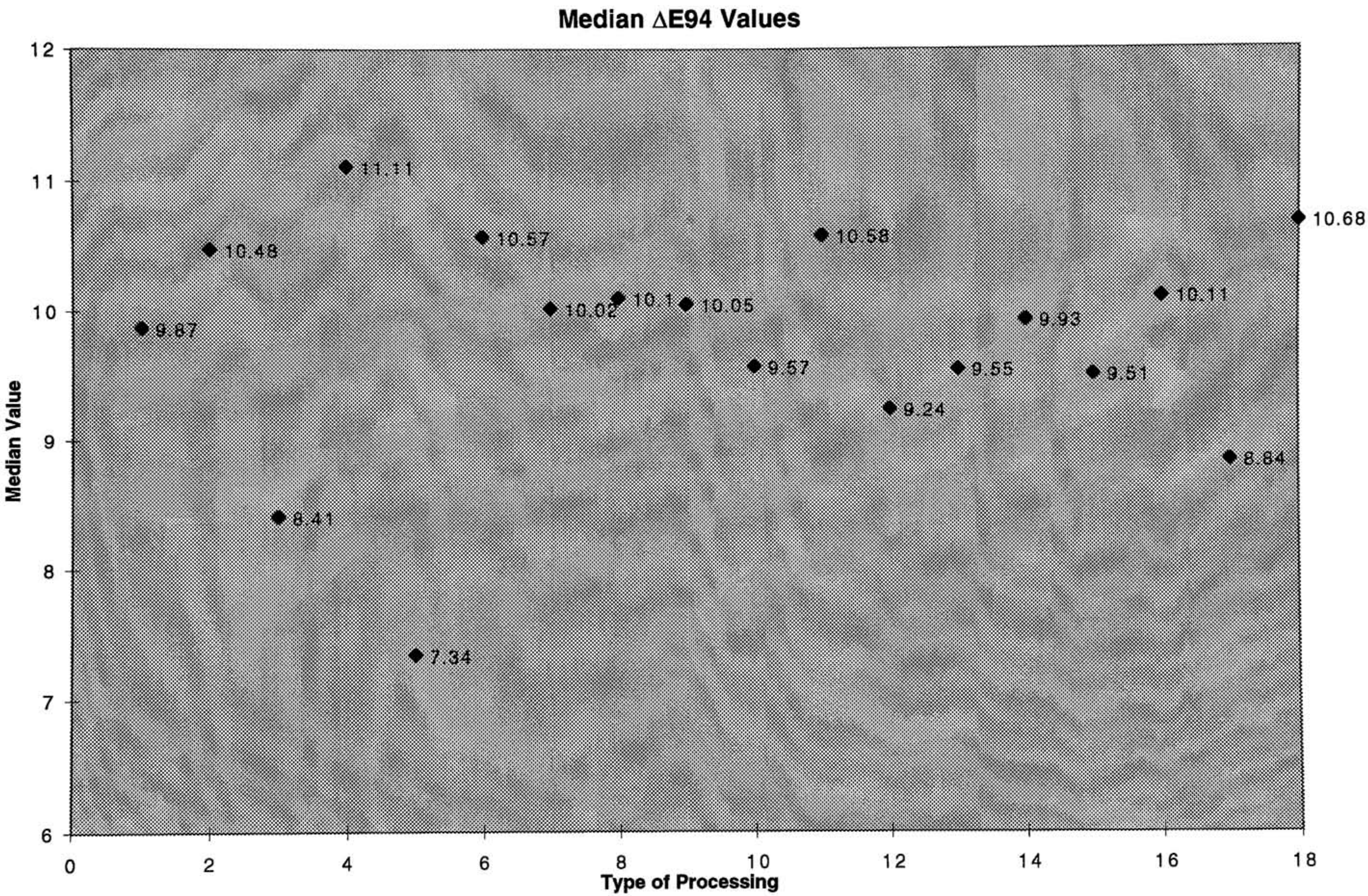
Figure 24.

Figure 25.



Median ΔE94 Values
AB1.1d
AB1.1m
AB1.3d
AB1.3m
AB1.5d
AB1.5m
L1.1d
L1.1m
L1.3d
L1.3m
L1.5d
L1.5m
LAB1.1d
LAB1.1m
LAB1.3d
LAB1.3m
LAB1.5d
LAB1.5m
Median Value
Type of Processing
0
2
4
6
8
10
12
14
16
18
6
7
8
9
10
11
12

Figure 26.


Median ΔE94 Values
Median Value
Type of Processing
12
11
10
9
8
7
6
0
2
4
6
8
10
12
14
16
18
9.87
10.48
8.41
11.11
7.34
10.57
10.02
10.1
10.05
9.57
10.58
9.24
9.55
9.93
9.51
10.11
8.84
10.68

Appendix 4.

Here are the results of the classification of the images used in the experiment. For each image, two sets of results are shown, one when the observer looked at the prints and no comparison with an original was perform, and the second set, when a CRT original was available.

For each image the frequency matrix and the final classification, in decreasing order from the most preferred, is given.

For image **“Building”** - no original.

Table 1.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	5	6	17	6	0
L1.1d	3	9	12	8	2
L1.3m	2	7	8	7	10
L1.3d	0	4	11	14	5
L1.5m	1	3	7	12	11
L1.5d	0	2	3	12	17
AB1.1m	0	2	12	18	2
AB1.1d	0	2	11	17	4
AB1.3m	0	4	18	10	2
AB1.3d	0	1	4	16	13
AB1.5m	2	6	10	15	1
AB1.5d	0	0	3	8	23
LAB1.1m	5	8	11	10	0
LAB1.1d	2	2	19	9	2
LAB1.3m	3	4	11	9	7
LAB1.3d	1	3	4	16	10
LAB1.5m	0	2	4	7	21
LAB1.5d	0	2	3	5	24
ORIGINAL	2	3	16	10	3

Table 2.

STIMULUS	Median	StdDev	Normalized median
L1.1m	1.028	0.69	-1.343
LAB1.1m	1.039	1.13	-1.323
L1.1d	1.089	1.04	-1.222
LAB1.1d	1.344	1.24	-0.716
AB1.3m	1.38	0.74	-0.644
ORIGINAL	1.406	1	-0.591
AB1.5m	1.549	0.77	-0.307
L1.3m	1.644	1.73	-0.117
LAB1.3m	1.667	1.16	-0.073
L1.3d	1.798	0.90	0.187
AB1.1m	1.823	0.65	0.238
AB1.1d	1.897	0.74	0.384
L1.5m	2.180	1.23	0.948
LAB1.3d	2.247	0.85	1.082
AB1.3d	2.448	0.92	1.482
L1.5d	2.716	1.69	2.015
LAB1.5m			
LAB1.5d			
AB1.5d			

For image **"Building"** - with original

Table 3.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	2	3	11	12	6
L1.1d	3	12	11	8	0
L1.3m	0	1	5	12	16
L1.3d	5	6	15	6	2
L1.5m	0	0	0	10	24
L1.5d	2	7	10	5	10
AB1.1m	1	10	16	7	0
AB1.1d	3	4	18	6	3
AB1.3m	4	11	12	6	1
AB1.3d	2	2	10	17	3
AB1.5m	5	8	12	7	2
AB1.5d	3	0	4	14	13
LAB1.1m	1	4	7	15	7
LAB1.1d	3	8	17	6	0
LAB1.3m	0	1	4	10	19
LAB1.3d	4	6	12	10	2
LAB1.5m	0	0	1	7	26
LAB1.5d	0	1	4	13	16
ORIGINAL	4	10	16	4	0

Table 4.

STIMULUS	Median	StdDev	Normalized median
AB1.3m	-0.332	0.98	-1.248
L1.1d	-0.317	1.15	-1.210
ORIGINAL	-0.312	0.71	-1.196
AB1.5m	-0.166	1.07	-0.833
LAB1.1d	-0.147	0.72	-0.784
AB1.1m	-0.125	0.78	-0.729
L1.3d	-0.099	0.99	-0.667
AB1.1d	0.055	0.92	-0.2805
LAB1.3d	0.083	0.95	-0.211
L1.5d	0.299	1.71	0.327
L1.1m	0.518	1.01	0.874
AB1.3d	0.541	0.79	0.930
LAB1.1m	0.578	1.07	1.0242
AB1.5d	0.67	1.21	1.252
L1.3m	0.718	1.02	1.373
LAB1.5d	0.720	1.02	1.377
LAB1.3m			
LAB1.5m			
L1.5m			

For image "**Fruits**" - no original

Table 5.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	7	12	12	3	0
L1.1d	2	5	13	11	3
L1.3m	2	8	14	8	2
L1.3d	0	1	4	10	19
L1.5m	0	3	4	11	16
L1.5d	0	0	0	9	25
AB1.1m	5	10	13	5	1
AB1.1d	2	5	14	11	2
AB1.3m	6	13	8	7	0
AB1.3d	2	7	9	11	5
AB1.5m	7	7	15	5	0
AB1.5d	1	1	7	15	10
LAB1.1m	4	12	11	6	1
LAB1.1d	1	9	10	11	3
LAB1.3m	3	11	7	8	5
LAB1.3d	0	0	2	9	23
LAB1.5m	0	2	2	8	22
LAB1.5d	0	0	0	3	31
ORIGINAL	4	7	13	7	3

Table 6.

STIMULUS	Median	StDev	Normalized median
L1.1m	0.746	0.73	-1.201
AB1.3m	0.757	0.86	-1.179
LAB1.1m	0.975	0.89	-0.782
AB1.1m	1.03	1.31	-0.681
AB1.5m	1.070	0.92	-0.607
ORIGINAL	1.205	0.78	-0.362
LAB1.3m	1.271	1.30	-0.241
L1.3m	1.334	1.11	-0.126
LAB1.1d	1.509	0.69	0.195
AB1.1d	1.522	0.89	0.218
L1.1d	1.57	0.95	0.306
AB1.3d	1.675	0.98	0.498
AB1.5d	2.301	1.47	1.642
L1.5m	2.674	1.04	2.323
L1.3d			
LAB1.5m			
LAB1.3d			
L1.5d			
LAB1.5d			

For image **"Fruits"** - with original

Table 7.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	3	7	15	8	1
L1.1d	7	15	10	1	1
L1.3m	2	2	3	19	8
L1.3d	8	8	4	10	4
L1.5m	0	0	1	7	26
L1.5d	0	1	7	11	15
AB1.1m	4	13	14	3	0
AB1.1d	0	2	9	18	5
AB1.3m	3	10	16	2	3
AB1.3d	3	3	21	7	0
AB1.5m	3	8	16	7	0
AB1.5d	2	5	8	15	4
LAB1.1m	5	5	14	9	1
LAB1.1d	3	18	10	3	0
LAB1.3m	0	1	4	20	9
LAB1.3d	0	4	10	6	14
LAB1.5m	0	0	0	5	29
LAB1.5d	0	1	1	8	24
ORIGINAL	2	6	19	7	0

Table 8.

STIMULUS	Median	StdDev	Normalized median
L1.1d	0.574	1.38	-1.451
LAB1.1d	0.669	0.98	-1.325
AB1.1m	0.861	0.77	-1.071
L1.3d	1.118	1.65	-0.731
AB1.3m	1.118	1.26	-0.731
AB1.5m	1.247	0.81	-0.560
L1.1m	1.342	0.86	-0.435
ORIGINAL	1.350	0.67	-0.426
LAB1.1m	1.376	0.86	-0.389
AB1.3d	1.400	0.59	-0.358
AB1.5d	2.031	1.03	0.477
AB1.1d	2.241	0.79	0.755
LAB1.3d	2.416	1.48	0.987
L1.3m	2.444	1.09	1.023
LAB1.3m	2.521	0.82	1.126
L1.5d	2.750	1.02	1.429
LAB1.5d	2.941	1.54	1.682
L1.5m			
LAB1.5m			

For image **"Landscape"** - no original

Table 9.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	0	7	15	12	0
L1.1d	1	5	10	14	4
L1.3m	6	9	10	7	2
L1.3d	1	3	8	13	9
L1.5m	3	11	7	6	7
L1.5d	2	3	12	9	8
AB1.1m	1	7	20	6	0
AB1.1d	0	5	13	13	3
AB1.3m	1	7	20	6	0
AB1.3d	0	1	10	13	10
AB1.5m	5	8	13	8	0
AB1.5d	1	0	5	16	12
LAB1.1m	4	14	9	6	1
LAB1.1d	4	8	15	7	0
LAB1.3m	1	10	13	4	6
LAB1.3d	0	4	13	12	5
LAB1.5m	1	8	8	7	10
LAB1.5d	0	1	8	16	9
ORIGINAL	3	7	14	9	1

Table 10.

STIMULUS	Median	StdDev	Normalized median
LAB1.1m	0.845	1.12	-1.72
L1.3m	1.121	1.19	-1.18
AB1.5m	1.234	1.03	-0.95
LAB1.1d	1.261	0.88	-0.90
L1.5m	1.361	1.95	-0.70
AB1.1m	1.384	0.64	-0.66
AB1.3m	1.384	0.64	-0.66
LAB1.3m	1.396	1.46	-0.64
ORIGINAL	1.437	0.84	-0.56
L1.1m	1.618	0.90	-0.20
AB1.1d	1.882	0.85	0.32
L1.5d	1.963	1.14	0.48
LAB1.3d	1.963	0.91	0.48
LAB1.5m	1.963	1.73	0.48
L1.1d	2.033	0.96	0.61
L1.3d	2.339	1.12	1.22
AB1.3d	2.414	0.83	1.36
LAB1.5d	2.452	0.81	1.44
AB1.5d	2.635	0.89	1.79

For image **"Landscape"** - with original

Table 11.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	0	4	17	11	2
L1.1d	3	6	15	6	4
L1.3m	0	2	8	16	8
L1.3d	0	7	17	9	1
L1.5m	0	0	3	18	13
L1.5d	1	9	13	9	2
AB1.1m	1	7	17	8	1
AB1.1d	0	4	18	12	0
AB1.3m	0	2	19	11	2
AB1.3d	4	5	14	9	2
AB1.5m	1	6	14	11	2
AB1.5d	3	3	11	9	8
LAB1.1m	0	2	8	15	9
LAB1.1d	0	4	14	12	4
LAB1.3m	1	1	3	17	12
LAB1.3d	2	7	16	7	2
LAB1.5m	0	0	1	8	25
LAB1.5d	1	8	15	7	3
ORIGINAL	0	2	15	14	3

Table 12.

STIMULUS	Median	StdDev	Normalized median
LAB1.3d	0.529	1.08	-0.934
AB1.1m	0.560	0.91	-0.854
L1.1d	0.564	1.31	-0.843
LAB1.5d	0.564	1.20	-0.843
L1.5d	0.570	1.13	-0.829
AB1.3d	0.605	1.08	-0.740
L1.3d	0.623	0.88	-0.694
AB1.5m	0.756	1.01	-0.352
AB1.1d	0.764	0.78	-0.331
L1.1m	0.809	0.87	-0.215
AB1.3m	0.836	0.76	-0.148
LAB1.1d	0.983	1.01	0.228
AB1.5d	1.059	1.44	0.422
ORIGINAL	1.059	0.81	0.422
L1.3m	1.064	1.04	0.435
LAB1.1m	1.135	1.09	0.617
LAB1.3m	1.717	1.23	2.107
L1.5m	1.892	0.71	2.556
LAB1.5m			

For image "**Musicians**" - no original

Table 13.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	3	5	9	10	7
L1.1d	2	9	15	8	0
L1.3m	4	5	9	11	5
L1.3d	0	0	2	16	16
L1.5m	0	0	3	7	24
L1.5d	0	0	0	5	29
AB1.1m	4	6	14	9	1
AB1.1d	3	5	17	6	3
AB1.3m	2	3	18	10	1
AB1.3d	3	4	9	14	4
AB1.5m	1	6	11	10	6
AB1.5d	1	2	10	16	5
LAB1.1m	4	12	10	6	2
LAB1.1d	1	5	11	12	5
LAB1.3m	1	5	11	10	7
LAB1.3d	0	0	1	9	24
LAB1.5m	0	1	1	9	23
LAB1.5d	0	0	0	6	28
ORIGINAL	1	3	8	14	8

Table 14.

STIMULUS	Median	StdDev	Normalized median
LAB1.1m	0.863	1.24	-2.17
L1.1d	1.157	0.83	-1.223
AB1.1m	1.255	0.83	-0.904
AB1.1d	1.283	0.98	-0.811
AB1.3m	1.417	0.69	-0.373
L1.3m	1.635	1.20	0.335
AB1.5m	1.654	1.16	0.399
L1.1m	1.743	1.31	0.689
LAB1.1d	1.743	1.02	0.689
LAB1.3m	1.743	1.16	0.689
AB1.3d	1.763	1.01	0.755
AB1.5d	1.814	0.84	0.919
ORIGINAL	1.844	1.06	1.017
LAB1.5m			
L1.3d			
L1.5m			
LAB1.3d			
LAB1.5d			
L1.5d			

For image **"Musicians"** - with original

Table 15.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	0	1	6	14	13
L1.1d	1	7	17	5	4
L1.3m	0	0	4	17	13
L1.3d	1	6	9	8	10
L1.5m	0	0	0	5	29
L1.5d	0	1	1	12	20
AB1.1m	2	8	14	8	2
AB1.1d	4	6	14	9	1
AB1.3m	2	4	5	18	5
AB1.3d	1	5	17	9	2
AB1.5m	1	1	6	13	13
AB1.5d	0	4	9	15	6
LAB1.1m	0	6	13	14	1
LAB1.1d	4	9	16	4	1
LAB1.3m	0	0	1	18	15
LAB1.3d	0	4	5	13	12
LAB1.5m	0	0	0	5	29
LAB1.5d	0	0	2	11	21
ORIGINAL	3	10	14	5	2

Table 16.

STIMULUS	Median	StdDev	Normalized median
LAB1.1d	1.101	0.98	-1.406
ORIGINAL	1.135	1.15	-1.357
AB1.1m	1.339	1.02	-1.059
AB1.1d	1.339	0.87	-1.059
L1.1d	1.367	1.13	-1.018
AB1.3d	1.479	0.86	-0.855
LAB1.1m	1.669	0.75	-0.578
L1.3d	1.963	1.57	-0.150
AB1.5d	2.13	1.01	0.092
AB1.3m	2.208	1.06	0.206
LAB1.3d	2.540	1.35	0.689
AB1.5m	2.630	1.14	0.821
L1.1m	2.656	0.98	0.859
L1.3m	2.715	0.79	0.945
LAB1.3m	2.861	0.58	1.158
L1.5d	2.992	1.26	1.348
LAB1.5d	2.992	0.93	1.348
LAB1.5m			
L1.5m			

For image **“Seed”** - no original

Table 17.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	6	10	12	5	1
L1.1d	4	12	11	7	0
L1.3m	5	9	10	9	1
L1.3d	1	7	14	7	5
L1.5m	3	7	10	8	6
L1.5d	2	3	13	12	4
AB1.1m	5	10	16	2	1
AB1.1d	2	11	14	6	1
AB1.3m	4	11	7	8	4
AB1.3d	0	2	14	11	7
AB1.5m	3	14	7	7	3
AB1.5d	0	2	11	9	12
LAB1.1m	5	16	6	6	1
LAB1.1d	2	9	12	7	4
LAB1.3m	6	7	7	7	7
LAB1.3d	0	0	10	12	12
LAB1.5m	2	5	8	11	8
LAB1.5d	0	0	0	16	18
ORIGINAL	0	1	11	13	9

Table 18.

STIMULUS	Median	StdDev	Normalized median
LAB1.1m	0.732	0.97	-1.607
AB1.5m	0.977	0.95	-1.093
L1.1m	1.056	0.98	-0.925
L1.1d	1.063	0.90	-0.910
AB1.1m	1.096	0.76	-0.842
AB1.1d	1.249	0.76	-0.519
AB1.3m	1.249	1.16	-0.519
L1.3m	1.263	1.14	-0.491
LAB1.1d	1.454	0.90	-0.089
LAB1.3m	1.522	1.58	0.054
L1.3d	1.590	0.81	0.198
L1.5m	1.645	1.16	0.313
L1.5d	1.858	1.12	0.761
AB1.3d	2.010	0.58	1.081
LAB1.5m	2.089	1.28	1.247
ORIGINAL	2.265	0.57	1.616
AB1.5d	2.317	0.68	1.726
LAB1.3d			
LAB1.5d			

For image **"Seed"** - with original

Table 19.

STIMULUS	IDEAL	EXCELLENT	GOOD	FAIR	POOR
L1.1m	0	0	5	15	14
L1.1d	0	5	7	18	4
L1.3m	0	1	3	15	15
L1.3d	5	3	14	10	2
L1.5m	0	0	0	9	25
L1.5d	2	7	18	7	0
AB1.1m	1	2	4	18	9
AB1.1d	0	3	10	16	5
AB1.3m	0	0	3	13	18
AB1.3d	1	10	18	5	0
AB1.5m	0	0	2	11	21
AB1.5d	3	15	10	4	2
LAB1.1m	0	0	4	14	16
LAB1.1d	3	4	15	11	1
LAB1.3m	0	0	3	4	27
LAB1.3d	4	13	12	4	1
LAB1.5m	0	0	0	4	30
LAB1.5d	3	7	8	12	4
ORIGINAL	3	10	15	5	1

Table 20.

STIMULUS	Median	StDev	Normalized median
AB1.5d	0.854	0.91	-1.468
LAB1.3d	0.915	0.96	-1.386
ORIGINAL	1.139	1.05	-1.085
AB1.3d	1.195	0.94	-1.010
L1.5d	1.288	0.68	-0.885
L1.3d	1.454	1.08	-0.661
LAB1.1d	1.474	0.89	-0.635
LAB1.5d	1.649	0.91	-0.400
AB1.1d	2.029	0.66	0.110
L1.1d	2.060	0.92	0.151
AB1.1m	2.366	1.46	0.561
L1.1m	2.635	0.92	0.922
L1.3m	2.708	0.79	1.021
LAB1.1m	2.776	2.19	1.112
AB1.3m	2.855	1.15	1.218
AB1.5m	2.855	1.24	1.218
LAB1.3m	2.855	0.99	1.218
L1.5m			
LAB1.5m			

References

1. Yule, J.A.C., *Principles of Color Reproduction,*Wiley: New York, 1967.

2. Cowan,W.B., and von Grunau,M.W., *Image Sameness: Some Experiments on Contrast in Simple Scenes,*Topical Meeting on Color Appearance, June 29-30, 1987, Annapolis, Maryland.

3. Evans,R.M.,Hanson,W.T.,and Brewer,W.L.,*Principles of Color Photography,* Wiley: New York. Chapman and Hall: London, 1953.

4. B.H.Marshall, and J.B.Guilford, *The Dependence of Hue, Tint, and Chroma Upon Area,* American Journal of Psychology, 46, 465-469 (1934).

5. M.F.Washburn, K.G.McLean, and A.Dodge, *The Effect of Area on the Pleasantness and Unpleasantness of Colors,* American Journal of Psychology, 46, 638-640 (1934).

6. R.W.Burnham, *The Dependence of Color upon Area,* American Psychologist, 4, 230-231 (1949).

7. E.Walton, and B.M.Morrison, *A Preliminary Study of the Affective Value of Colored Lights*, Journal of Applied Psychology, 15, 294-303 (1931).

8. J.F.Dashiell, *Children's Sense of Harmonies in Colors and Tones*, Journal of Experimental Psychology, 2, 466-475 (1917).

9. W.E.Walton, *The Sensitivity of Children and Adults to Color Harmony*, Psychological Monographs, 45, 51-62 (1933).

10. Committee on Colorimetry Optical Society of America, *The Science of Color*, New York: Thomas Y. Crowell Company, 1953.

11. E.C.Allen, and J.P.Guilford, *Factors Determining the Affective Values of Color Combinations*, American Journal of Psychology, 48, 643-648 (1936).

12. M.F.Washburn, D.Haight, and J.Regensburg, *The Relation of Pleasantness of Color Combinations to that of the Colors Seen Singly*, American Journal of Psychology, 32, 145-146 (1921).

13. P.Moon, and D.E.Spencer, *Area in Color Harmony*, Journal of Optical Society of America, 34, 93-103 (1944).

14. N.C.Meier, *Art in Human Affairs: an Introduction to the Psychology of Art*, McGraw-Hill, 1942.

15. Gordon E.Legge, David H.Parish, Andrew Luebker, and Lee H.Wurm, *Psychophysics of reading. XI. Comparing color contrast and luminance contrast*, J.Opt.Soc.Am. A/Vol. 7, No. 10/October 1990.

16. McIlhagga,W.H.,Mullen,K.T., *Contour Integration with Colour and Luminance Contrast*,Vision Res., Vol. 36, No. 9, pp.1265-1279, 1996.

17. Tai-Lioan Chen, and Chi-Yuang Yu, *The Relationship between Visual Acuity and Color Contrast in the OSA Uniform Color Space*,Color research and application, Vol. 21, No. 1, February 1996.

18. Wyszecki, G., and Stiles, W.S., *Color Science: Concepts and Methods. Quantitative Data and Formulae*, 2nd Ed., Wiley, New York, 1982.

19. Alpern, M., *Relation between Brightness and Color Contrast,* J.Opt.Soc.Am., Vol.54,No.12,p.1491-1492,Dec.1964.

20. Fairchild, M.D., and Berns, R.S.,*Image Color-Appearence Specification Through Extension of CIELAB,* Color research and application, Vol. 18, No. 3, June 1993.

21. Fairchild, M.D., *Formulation and testing of an Incomplete-Chromatic-Adaptation Model*, Color research and application, Vol. 16, No. 4, August 1991.

22. Lucassen, M., *Quantitative Studies of Color Constancy*, 1993.

23. Gescheider, G.A., *Psychophysics, Method, Theory, and Application,* Lawrence Erlabaum Associates, Publishers, !985, Hillsdale, New Jersey.

24. Buser, P., and Imbert, M., *Vision,* The MIT Press Cambridge, Massachusetts, London, England, 1992.

25. Graham, N.V.S. *Visual Pattern Analysers*, Oxford Science Publications, !989.

15. Inter-Society Color Council Proceeding 1971, *Optimum Reproduction of Color*, Williamsburg, Virginia, Jan. 31 - Feb. 3, 1971.

26. Fairchild, M.D., *Considering the Surround in Device-Independent Color Imaging*, Color research and application, Vol. 20, No. 6, Decmber 1995.

27. *Testing Colour-Appearance Models: Guidelines for Coordinated Research,* Color research and application, Vol. 20, No. 4, August 1995.

28. Sachs, L., *Applied Statistics, A Handbook of Techniques,* Second Edition, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

29. Fairchild, M.D., Berns, R.S., *Image Color-Appearance Specification Through Extension of CIELAB*, Color research and application, Vol.18, Number 3, June 1993.

30. Berns, R.S., *Methods for characterizing CRT displays*, Displays, Volume 16, Number 4, 1996.

31. Wright, W.D, *The Measurement of Colour*, 4th edition, Van Nostrand Reinhold Company, 1969.

32. Committee on Colorimetry - Optical Society of America, *The Science of Color*, New York: Thomas Y. Crowell Company, 1953.

33. Cowan, W.B., *The Computational Approach to Colour Reproduction,* in Topical Meeting on Color Appearance, June 29-30, 1987, Annapolis, Maryland.